

Uncovering the structure of agreeableness from self-report measures

Michael L. Crowe¹ | Donald R. Lynam² | Joshua D. Miller¹

¹University of Georgia

²Purdue University

Correspondence

Joshua D. Miller, 125 Baldwin Street,
Department of Psychology, University of
Georgia, Athens, GA 30602

Email: jdmiller@uga.edu

Abstract

Objective: Although there are several models of the lower-order structure of Agreeableness, empirically derived descriptions of this domain are largely nonexistent. We examined the factor structure of Agreeableness items from multiple scales in order to empirically determine the facet-level structure of the domain.

Method: Participants ($N = 1,205$; 73% female; 84% White; $M_{age} = 35.5$, $SD = 17.26$) completed 131 items from 22 scales measuring Agreeableness.

Results: A series of factor analyses was conducted on 104 items to identify factor emergence of the domain from a single factor to increasingly more specific factors. A five-factor solution consisting of facets labeled *Compassion*, *Morality*, *Trust*, *Affability*, and *Modesty* was identified as most appropriate. Factors from all levels of the construct were compared to current measures of the domain as well as a number of criterion variables. The patterns of association with criterion variables at the lower level of the Agreeableness domain showed significant divergence.

Conclusions: The current results highlight how specific Agreeableness traits unfold from broader to more specific facets and how these traits are represented in existing measures of this important domain.

KEY WORDS

Big Five, Five-Factor Model, personality structure

1 | INTRODUCTION

The development of general trait taxonomies represented a major turning point in personality science, as they provided models to systematically organize the breadth of findings in the personality field (John & Srivastava, 1999). By far, the most widely known personality taxonomy is the Big Five (B5)/Five-Factor Model (FFM), which organizes personality into five primary, higher-order domains: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. Although the B5/FFM has been invaluable for the progression of personality literature, there are still aspects of the model that require further development. Years of empirical examination have yielded broad consensus at the domain level of the model (Digman, 1990; John, Naumann, & Soto, 2008; John & Srivastava, 1999), but

research focusing solely on personality domains can be problematic, as the heterogeneity of the domains may make interpretation and theory development difficult (Smith, McCarthy, & Zapolski, 2009). Underlying facets have been developed for these broad domains within various models, but the facets have typically been rationally identified. In order for the field to continue to progress, it is important to develop empirically derived, lower-order structures for these constructs. For instance, Watson, Stasik, Ellickson-Larew, and Stanton (2015) used a number of Extraversion subscales to identify empirically four facets of the broader domain that manifested somewhat divergent relations with psychopathological outcomes. Similarly, Roberts, Chernyshenko, Stark, and Goldberg (2005) empirically identified six facets of the Conscientiousness domain. The goal of the present study is to continue this process for Agreeableness by using a

bottom-up approach to develop a comprehensive lower-order structure based on an item-level examination of existing scales.

Agreeableness (vs. Antagonism) is an interpersonal dimension of personality reflecting individual differences in cooperation and the motivation to maintain positive social relations. It has been associated with a range of outcomes, including management of interpersonal conflict, prejudice, and peer relationships (Gleason, Jensen-Campbell, & South Richardson, 2004; Jensen-Campbell et al., 2002; Jensen-Campbell & Graziano, 2001; Sibley & Duckitt, 2008). Agreeableness has been linked to job performance (Witt, Burke, Barrick, & Mount, 2002), pro-social behavior, volunteering (Carlo, Okun, Knight, & De Guzman, 2005; Graziano, Habashi, Sheese, & Tobin, 2007), and positive affect (DeNeve & Cooper, 1998). At the opposite pole, Antagonism is an important correlate of crime and aggression (Jones, Miller, & Lynam, 2011; Miller & Lynam, 2001) and a central component of problematic personality disorders (PDs) such as psychopathy, antisocial PD, and narcissistic PD (e.g., Lynam & Widiger, 2001; Miller, Lynam, Widiger, & Leukefeld, 2001; Samuel & Widiger, 2008). Although Agreeableness has proven to be a valuable construct, analysis at the domain level alone provides only a limited understanding of its relationship to outcomes of interest, as the strength of these findings may vary across the lower-order scales (Smith et al., 2009). For example, in the Jones et al. (2011) meta-analysis of the FFM correlates of antisocial behavior, the Agreeableness facets of Straightforwardness and Compliance were stronger correlates than Trust and Tendermindedness.

1.1 | Current models (and measures) of agreeableness

Much of the research on the structure Agreeableness has focused on a two-level model, in which a single domain, representing the broadest conceptualization of the latent trait, is underlaid by a second, lower level composed of four or more facets. One only needs to compare the facet structures of the major Agreeableness models to see that there is not a consensual lower-order structure.

The Big Five Inventory (BFI; John, Donahue, & Kentle, 1991) is perhaps the most popular measure of the B5 and was developed using the lexical approach. Two facets of Agreeableness (Altruism and Compliance) have been identified within the BFI (Soto & John, 2009). The items in these facets were selected to make the scales congruent with the NEO-PI-R facets of the same name. Lexical analyses have been used in an attempt to identify additional facets of the Big Five (Perugini & Gallucci, 1997; Saucier & Ostendorf, 1999). Saucier and Ostendorf (1999) identified four facets of Agreeableness, including Warmth-Affection (e.g.,

affectionate, sentimental, compassionate vs. cold and unsympathetic), Gentleness (e.g., agreeable, amiable vs. harsh, antagonistic, combative), Generosity (e.g., charitable, helpful, generous vs. greedy, selfish), and Modesty-Humility (e.g., modest vs. egotistical, boastful, vain). Perugini and Gallucci (1997) identified five Agreeableness facets: Sympathy (e.g., tolerant, conciliatory, sympathetic), Tendermindedness (e.g., sentimental, romantic, affectionate), Friendliness (e.g., loyal, honest, faithful), Overbearance (e.g., overbearing, authoritarian, tyrannical), and Hostility (e.g., revengeful, hostile, pitiless). In general, however, these facets are seldom used in research employing the BFI.

Agreeableness, as measured by the predominant measure of the FFM, the Revised NEO Personality Inventory (NEO-PI-R; Costa & McCrae, 1992), is a two-level model that includes six facets. NEO-PI-R Agreeableness facets were identified not through an empirical analysis, but through an examination of previous literature (Costa, McCrae, & Dye, 1991). The facets were specified as follows: Trust (belief in the sincerity and good intentions of others), Straightforwardness (sincerity and unwillingness to manipulate others), Altruism (concern for the welfare of others), Compliance (willingness to cooperate with others, ability to inhibit aggression and forgive others when faced with a potential conflict), Modesty (humility), and Tendermindedness (attitude of sympathy for others).

The HEXACO (Lee & Ashton, 2004) is a more recently developed personality model and related assessment inventory. Like the NEO-PI-R, the HEXACO represents a two-level conceptualization of Agreeableness. The HEXACO originated from more recent lexical studies in which six (rather than five) lexical factors emerge (Ashton et al., 2004; Saucier, 2009). In the HEXACO, the variance of the Agreeableness domain (as defined by B5/FFM) is spread across Honesty/Humility and Agreeableness factors (see Ashton, Lee, & De Vries, 2014, for a more thorough review). Facets of these domains were identified through an examination of item content obtained in lexical studies (Lee & Ashton, 2004). The facets of Honesty/Humility were specified as Sincerity (interpersonally genuine, unwillingness to manipulate others), Fairness (unwillingness to cheat or take advantage of others to get ahead), Greed Avoidance (uninterested in or unmotivated by possession of wealth or status symbols), and Modesty (unassuming, no desire for special treatment). The facets of Agreeableness were identified as Forgiveness (trust and liking toward others, even after mistreatment), Gentleness (interpersonal lenience or reluctance to judge others harshly), Flexibility (willingness to compromise and cooperate), and Patience (interpersonally imperturbable or high threshold for anger). It is noteworthy that HEXACO Agreeableness does not include altruism or a concern for the well-being of others, a facet that is primary to many other models of this domain (John et al., 2008). However, the

HEXACO-PI-R does include an Altruism “interstitial” scale intended to assess traits of sympathy and soft-heartedness (Lee & Ashton, 2006). This scale is not included in any one HEXACO domain, as lexical analyses suggest that the item content related to this factor tends to have significant associations with multiple domains (i.e., Honesty-Humility, Agreeableness, and Emotionality).

The Faceted Inventory of the Five-Factor Model (FI-FFM; E. E. Simms, 2009) offers another two-level model of Agreeableness. The FI-FFM identified facets of the FFM through a construct validation approach (L. J. Simms & Watson, 2007), and it has been successfully used in a few studies (Naragon-Gainey, Watson, & Markon, 2009; Watson et al., 2015; Watson, Stasik, Ro, & Clark, 2013), although it has been less extensively used and validated than other B5/FFM assessments. The FI-FFM identified the following facets of Agreeableness: Empathy (pro-social, pleasant, sensitive to the needs of others), Trust vs. Cynicism (belief in the goodness of others and willingness to trust), Straightforwardness vs. Manipulativeness (willingness to take advantage of or manipulate others to get ahead), and Modesty (modesty, humility).

These various conceptualizations of Agreeableness are difficult to integrate. There are clearly common elements across many of them (e.g., sympathy, morality, conflict avoidance), but some facets are unique as well (e.g., greed avoidance). More importantly, given concerns related to the jingle-jangle fallacy (e.g., Block, 1995), one cannot assume that similarly named facets measure the same construct, nor can one assume that facets with different labels do, in fact, measure substantively different constructs.

1.2 | Structure of agreeableness

Further empirical analysis is needed to identify the facets underlying this domain (Graziano & Tobin, 2017). There are potential shortcomings associated with limiting personality organization to only two levels of analysis (i.e., domains and facets). DeYoung, Quilty, and Peterson (2007) have shown that there is value in a level of analysis between facets and domains. Specifically, DeYoung and colleagues’ Big Five Aspect Scales represent the Agreeableness domain via two midrange factors: Compassion (i.e., interpersonal warmth, sympathy, tenderness) and Politeness (i.e., cooperativeness, compliance, and straightforwardness). A recent revision of the BFI—the BFI-2—also represents Agreeableness at a midrange of specificity with three factors: Compassion, Respectfulness, and Trust (Soto & John, 2017).

Goldberg (2006) offered a methodology allowing a “hierarchical” model of personality to be studied at a range of intermediate levels of facet specificity. This approach allows for exploration of this intermediate space while

accommodating and organizing each of the previously discussed facet-level models.

The goal of the present research is to describe meaningful levels of the Agreeableness domain using an adaptation of Goldberg’s (2006) “Bass-Ackwards” approach.¹ In this approach, factor analyses are used to extract from one to many factors while saving the factor scores from each analysis. The relations among the factor scores from adjoining levels reveal the unfolding structure of the domain. This method provides a map of factor emergence for each level of Agreeableness from broadest to most precise. Importantly, this method allows for the identification of the relations between factors at various levels of specificity, a better understanding of when (at which level) each of the facets emerge, and the centrality of each facet to the Agreeableness construct. In the current study, we modified Goldberg’s approach by allowing factors within a given level to be correlated rather than forcing them to be orthogonal. Unlike the B5/FFM domains, which are thought to index relatively distinct constructs, the Agreeableness facets all assess interrelated content; thus, although the assumption of orthogonal factors is reasonable in the case of the B5/FFM, it is not reasonable for a within-domain examination of Agreeableness facets. Although previous research on the lower-order structures of Conscientiousness and Extraversion involved factor-analyzing scale scores (Roberts et al., 2005; Watson et al., 2015), we chose instead to work at the item level, as concerns regarding item parceling in confirmatory factor analysis (Marsh, Lüdtke, Nagengast, Morin, & Von Davier, 2013) apply to the present analyses. Facet scales are parcels of items. According to Marsh et al., parcels are appropriate only under specific circumstances, namely, when the parcels are unidimensional and lack high cross-loadings. Given the diversity in construction and content of the various facet scales, these requirements are unlikely to be met. By conducting analyses at the item level, we maximize the flexibility of the potential factor structure and increase our ability to identify a stable lower-order structure.

2 | METHOD

2.1 | Participants and procedure

A total of 1,289 participants were recruited from Amazon’s Mechanical Turk (MTurk) website. Participants were required to be 18 years of age or older and to reside in the United States. Participants were paid \$1.00 for their participation. Of the 1,289 participants who completed informed consent, 81 participants were removed for failing one or both of the validity scales (see Measures section), and three participants were removed for failing to complete more than 50% of the items. Pairwise deletion was utilized for the remaining participants with missing data. The final dataset consisted of

1,205 participants (73% female; 84% White; $M_{age} = 35.5$, $SD = 17.26$).²

2.2 | Measures³

2.2.1 | Agreeableness items

Big Five Inventory (BFI)

The BFI (John et al., 1991; nine Agreeableness items) is a brief (44-item) measure of the Big Five personality domains. The alphas of the five domain scales ranged from .82 (Agreeableness) to .89 (Neuroticism).

Big Five Aspects Scale (BFAS)

The BFAS (DeYoung et al., 2007; 20 Agreeableness items) separates each of the five major personality domains into two aspects, which can be understood to fall structurally between the facet and domain levels of personality models. Only the Agreeableness items were collected for this study. These items are organized into aspects of Compassion ($\alpha = .91$) and Politeness ($\alpha = .79$). Nine of 10 Compassion items and all 10 Politeness items were used in the final pool of Agreeableness items.⁴

Faceted Inventory of the Five-Factor Model (FI-FFM)

The FI-FFM (E. E. Simms, 2009; 42 Agreeableness items) was developed through a construct validation approach. Only the Agreeableness items were collected for the present study. The FI-FFM Agreeableness domain contains four facets: Empathy ($\alpha = .87$), Trust ($\alpha = .91$), Straightforwardness ($\alpha = .83$), and Modesty ($\alpha = .85$). Nine of 10 Empathy items, seven of 10 Trust items, 10 of 10 Straightforwardness items, and nine of 10 Modesty items were used in the final pool of Agreeableness items.

HEXACO-PI-R

The HEXACO-PI-R (Lee & Ashton, 2006; 16 Agreeableness, 16 Honesty-Humility, and four Altruism items) is a 100-item measure of six different personality domains: Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience. Each of the domains can be separated into four facets. For the present study, the Honesty-Humility and Agreeableness items were administered along with the four-item Altruism interstitial scale. The alphas of the facet scales ranged from .72 (Modesty) to .83 (Greed Avoidance) for the Honesty-Humility domain and from .66 (Flexibility) to .77 (Forgiveness) for the Agreeableness domain. Alpha for the Altruism interstitial scale was .65. Three of four Altruism items were used in the final pool of Agreeableness items. From the Honesty-Humility domain, two of four Sincerity items, two of four Fairness

items, three of four Greed Avoidance items, and four of four Modesty items were used in the final item pool. From the Agreeableness domain, three of four Forgiveness, four of four Gentleness, three of four Flexibility, and three of four Patience items were used in the final pool of Agreeableness items.

International Personality Item Pool (IPIP)

The IPIP (Goldberg et al., 2006; 24 Agreeableness items) is a collection of publicly available items and scales. The 24 Agreeableness items from the 120-item IPIP-NEO-PI-R (Maples, Guan, Carter, & Miller, 2014) were collected for the present study. These 24 items can be divided into six different facet scales whose alphas ranged from .72 (both Modesty and Morality) to .88 (Trust). One of four Trust items, three of four Morality items, three of four Altruism items, two of four Cooperative items, one of four Modesty items, and four of four Sympathy items were used in the final pool of Agreeableness items.

2.2.2 | Criterion measures

Interpersonal Adjectives Scale (IAS)

The IAS (Wiggins, 1995) uses self-report ratings of 64 adjectives to provide scores relevant to the interpersonal circumplex (IPC) model of personality. The IAS can be used to generate eight separate octant scores. The alphas of these octants ranged from .78 (unassuming-ingenuous) to .92 (warm-agreeable). For the present analyses, only the scores representing the two axes of the IAS (i.e., love and dominance) were used.

Crime and Analogous Behavior scale (CAB)

The CAB (Miller & Lynam, 2003) assesses substance use and antisocial behavior. The substance use ($\alpha = .72$) variable was calculated by counting the number of five different substances participants endorsed trying. The antisocial behavior ($\alpha = .66$) variable was calculated by counting the number of nine different behaviors endorsed.

Reactive and Proactive Aggression Questionnaire (RPAQ)

The RPAQ (Raine et al., 2006) is a measure of aggression that includes both a total aggression score ($\alpha = .88$) as well as reactive aggression ($\alpha = .83$) and proactive aggression ($\alpha = .85$) subscale scores.

Validity scales

Two validity scales from the Elemental Psychopathy Assessment (Lynam et al., 2011) were used—the Infrequency Scale (e.g., “I try to eat something almost every day”; reversed) and the Too Good to Be True Scale (e.g., “I have never in my life been angry at another person”). Participants’ data

were omitted if they received a score of 4 or more on the Infrequency Scale or a score of 3 or more on the Too Good to Be True Scale.

2.3 | Data analysis

Before data collection occurred, all of the Agreeableness items from each of the relevant scales listed above were intermixed into a single scale of 131 items. Items not originally presented in the form of a complete sentence were put into sentence form for the sake of consistency (i.e., "I trust others" rather than simply "trust others"). All items were presented in a random order to participants, who rated their agreement with each item on the same 1 (*Strongly disagree*) to 5 (*Strongly agree*) scale.

After data collection was complete, all relevant items were correlated with one another in order to identify duplicate or excessively overlapping items so as to reduce the likelihood of extracting bloated specific factors. Twenty item pairs were identified with correlations greater than .65, and an item from each of the pairs was removed from the pool, yielding a total of 113 Agreeableness items.⁵ A principal-factors analysis was then conducted on this item pool to identify items that loaded poorly on the first unrotated factor. Items with factor loadings less than .30 on this general Agreeableness factor were removed for being unrepresentative of the general factor (Osborne, Costello, & Kellow, 2008). This process removed an additional nine items from the pool. Therefore, the final pool on which the following structural analysis was conducted consisted of 104 Agreeableness items.⁶

After the final item pool was created, the structure of Agreeableness was evaluated. All factor solutions were identified using the principal axis factoring method with promax rotation. A single unrotated factor was extracted; then rotated solutions of successively more factors were extracted until one of the factors was either too specific to be meaningful or was no longer interpretable. At each step in the process, the factor scores were saved so that different levels of the factor structures could be correlated and compared. The identified factors were correlated with existing Agreeableness scales and relevant external criterion variables (e.g., the remaining BFI domains).

3 | RESULTS

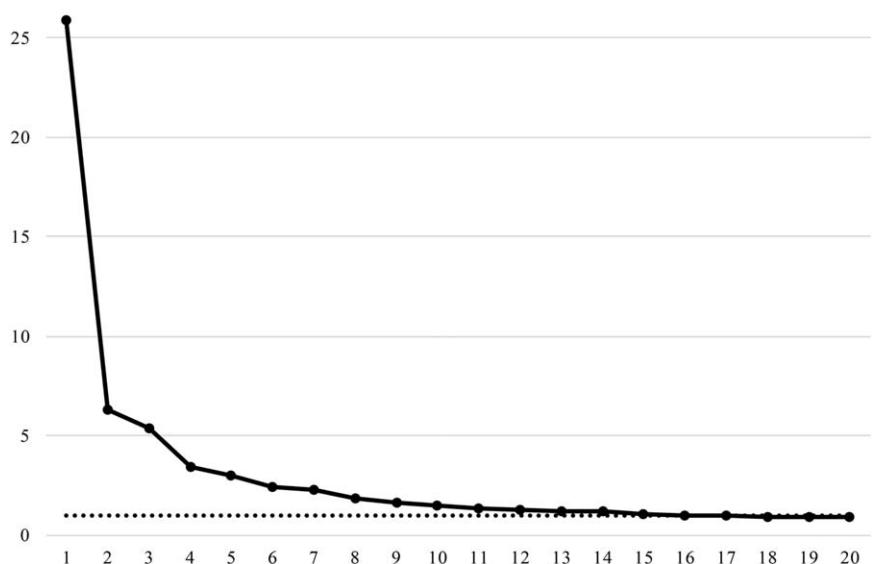
The first unrotated Agreeableness factor accounted for 24% of the total variance. The first 20 eigenvalues of this analysis are shown in the scree plot presented in Figure 1. Following the single-factor solution, a series of successively larger solutions was examined. Although the goal of this analysis was to characterize the relations between multiple structures of

Agreeableness rather than to identify the "ideal" number of factors, we employed several approaches to identifying the optimal number of factors. The eigenvalues were examined in the form of a scree plot (Figure 1), which suggested the extraction of four to seven factors. A parallel analysis (Horn, 1965) identified 10 factors. Velicer's minimum average partial (MAP) test was also considered (Velicer, 1976), which indicated the presence of eight factors. Analyses progressed until a nine-factor solution was identified. In the nine-factor solution, zero items had their highest loading on one of the nine factors, suggesting that eight factors are the maximum number of unique factors that should be considered for further evaluation. After evaluating each of the solutions, the five-factor model was selected as the final factor solution. The factors that emerged beyond this level were composed of overly specific content. Although the analyses were conducted through the eight-factor solution, our analyses will include only those results related to the one- through five-factor solutions (see the supplemental materials for data on the six- through eight-factor solutions).⁷

Given the goal of the present analyses, replicability of the analysis is of primary importance. As such, factor scores from the principal-factors solutions were compared to those derived with maximum likelihood (ML) estimation (see Saucier, 2003, for a similar method). The one-factor through five-factor solutions proved to be highly robust to different factor-analytic techniques, as each of the principal-axis factor scores correlated greater than .99 with corresponding factors derived through ML estimation. However, when six factors were extracted, the correspondence across extraction methods diminished, as ML estimation yielded two factors (6.4 and 6.5) whose score estimates correlated less than .95 (.92 and .93, respectively) with the principal-axis-derived scores.

In the interest of interpretability and parsimony, the content of each factor through the eight-factor solution was also evaluated. The content of each of the factors through the five-factor solution was generally interpretable; the factors were composed of items with generally high loadings and both positively and negatively keyed items. The six-factor solution introduced a number of issues that made it more difficult to interpret. In moving from the five-factor to six-factor solution, Compassion (6.1), Cooperation (6.2), Trust (6.3), and Morality (6.6) all emerged in forms generally consistent with their counterparts in the five-factor solution. Modesty (5.5), however, broke into two factors (Factors 6.4 and 6.5 in the supplementary materials) with content that was more difficult to interpret. Factor 6.5 remained a Modesty versus Grandiosity factor, but Factor 6.4 was problematic, appearing to be more of a social desirability factor. This

Agreeableness Scree Plot

**FIGURE 1** Agreeableness scree plot. Dotted line indicates an eigenvalue of 1

factor's highest-loading item loaded at .57 compared to the highest-loading items on the other factors, which ranged from .66 to .78 with a mean of .72. The number of items on the other factors with higher loadings than .57 ranged from 6 to 24 with a mean of 13. Finally, five of the 10 highest-loadings items on this factor had higher loadings on other factors. It was with all of these criteria in mind that the five-factor solution was selected as the final factor solution.

3.1 | Emergence of Agreeableness factors

Factor score correlations for each of the progressive solutions from one factor to five factors are presented in Figure 2, and example items for each of the solutions from one factor to five factors are presented in Table 2.⁸ To characterize each of the factors across the various solutions, correlations between the factor score and the existing Agreeableness

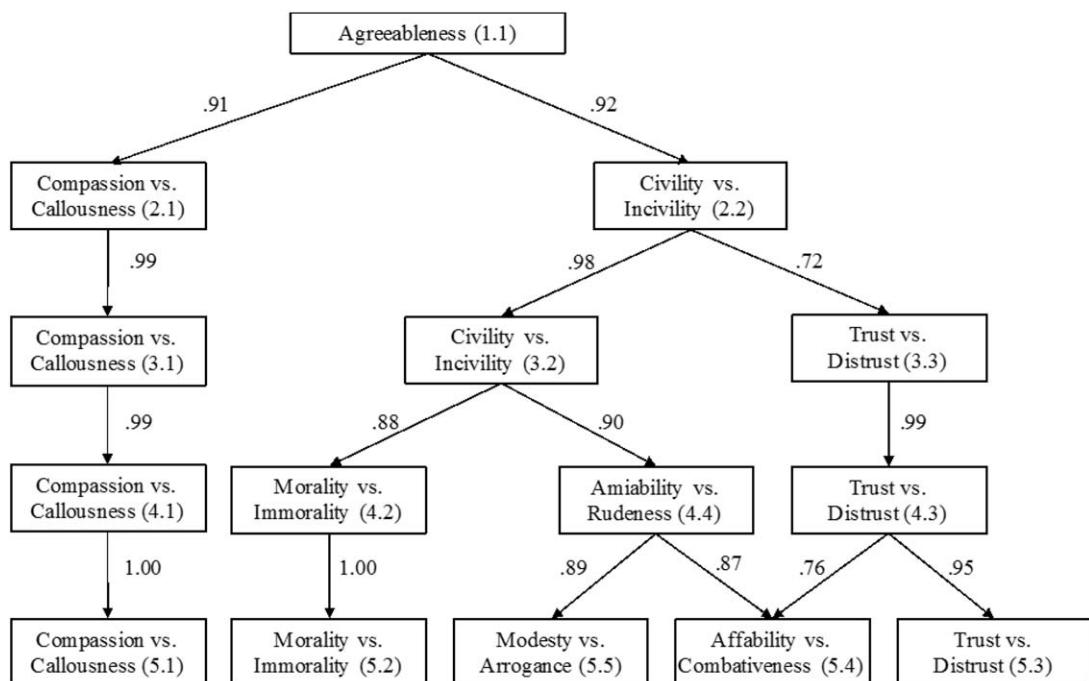
**FIGURE 2** Hierarchical structure of the agreeableness domain. Only correlations going from one level of the hierarchy to the next are identified. Correlations less than .70 were removed

TABLE 1 Agreeableness scales descriptive statistics

Scale	Items	n	α	M (SD)
BFI-A	9 (9)	1,203	.82	3.70 (0.63)
IPIP-NEO-A		1,194	.87	3.71 (0.49)
Trust	4 (1)	1,202	.88	3.24 (0.85)
Morality	4 (3)	1,204	.72	3.85 (0.73)
Altruism	4 (3)	1,201	.77	3.95 (0.66)
Cooperativeness	4 (2)	1,201	.74	3.95 (0.78)
Modesty	4 (1)	1,204	.72	3.49 (0.78)
Sympathy	4 (4)	1,200	.75	3.75 (0.74)
HEX-A		1,200	.86	3.06 (0.61)
Forgiveness	4 (3)	1,203	.77	2.63 (0.82)
Gentleness	4 (4)	1,202	.68	3.36 (0.74)
Flexibility	4 (3)	1,203	.66	3.01 (0.78)
Patience	4 (3)	1,203	.74	3.22 (0.83)
HEX-H		1,201	.85	3.54 (0.66)
Sincerity	4 (2)	1,204	.73	3.41 (0.89)
Fairness	4 (2)	1,202	.82	3.65 (1.04)
Greed Avoidance	4 (3)	1,202	.83	3.20 (1.00)
Modesty	4 (4)	1,205	.72	3.90 (0.74)
HEX-Altruism	4 (3)	1,204	.65	3.93 (0.68)
FI-FFM-A		1,194	.92	3.54 (0.49)
Empathy	10 (9)	1,201	.87	3.81 (0.63)
Trust	11 (7)	1,200	.91	3.09 (0.73)
Straightforwardness	11 (10)	1,201	.83	3.37 (0.70)
Modesty	10 (9)	1,201	.85	3.86 (0.61)
BFAS-A		1,199	.90	3.93 (0.54)
Compassion	10 (9)	1,200	.91	3.97 (0.65)
Politeness	10 (10)	1,202	.79	3.89 (0.57)
Total	131 (104)	1,179	0.97	3.59 (0.47)

Note. BFI-A = Big Five Inventory-Agreeableness; IPIP-NEO-A = IPIP-NEO-Agreeableness; HEX-A = HEXACO-Agreeableness; HEX-H = HEXACO-Honesty-Humility; HEX-Altruism = HEXACO-Altruism; FI-FFM-A = Faceted Inventory of the Five-Factor Model-Agreeableness; BFAS-A = Big Five Aspects Scale-Agreeableness. As described in the Data Analysis section, items that did not load strongly on the first unrotated factor were removed from the Agreeableness item pool. The item numbers in parentheses indicate the number from each scale that was included in the final item pool.

scales are presented in Table 3. The similarities of the correlational profiles for the factor scores are presented in Table 4; for example, the .89 at the top of the lower diagonal of Table 4 is the correlation between columns F1.1 and F2.1 from Table 3. Table 4 also provides the correlations of the factor scores with one another across the levels of the analyses; for example, the .91 at the top left-hand corner of the upper diagonal represents the correlation between the factor scores for F1.1 and F2.1. The first unrotated factor manifested strong correlations with all of the domain-level Agreeableness scales, with correlations ranging from .67

(HEXACO Honesty-Humility) to .94 (BFAS Agreeableness, FI-FFM Agreeableness). This factor also demonstrated substantial correlations with the narrower facets, with correlations ranging from .35 (IPIP-NEO Modesty) to .86 (FI-FFM Empathy), with a median of .65. This Agreeableness factor (Agreeableness, 1.1) correlated roughly equivalently with both factors of the two-factor solution ($r_s = .91$ and $.92$).

The first factor of the two-factor solution, Compassion (vs. Callousness; 2.1), was composed of items relating to a concern for the feelings of others. It manifested domain-level correlations that ranged from .45 (HEXACO Honesty-Humility) to .90 (BFAS Agreeableness) and facet-level correlations that ranged from .22 (IPIP-NEO Modesty) to .94 (HEXACO Empathy), with a median of .48. This factor had its highest correlations with scales assessing Altruism (IPIP-NEO, HEXACO), Sympathy (IPIP-NEO), Empathy (FI-FFM), and Compassion (BFAS), and (as shown in Table 4) its profile remained consistent from the two-factor solution through the five-factor solution, with profile similarities greater than .98 throughout. The second factor of the two-factor solution (Civility vs. Incivility; 2.2) was composed of more heterogeneous item content related to (at the negative pole) manipulativeness, dishonesty, and contentiousness. This factor manifested correlations with the domain-level scales that ranged from .70 (HEXACO Agreeableness) to .91 (FI-FFM Agreeableness) and with the facets that ranged from .41 (IPIP-NEO Trust) to .87 (BFAS Politeness), with a median of .59. Its highest facet-level correlations included Morality and Cooperativeness from the IPIP-NEO, Straightforwardness and Modesty from the FI-FFM, and Politeness from the BFAS.

The content of Compassion (2.1) and Civility (2.2) generally carried over to the first (Compassion, 3.1) and second factors (Civility, 3.2) of the three-factor solution.⁹ Compassion (3.1) manifested a profile similarity of .99 with its counterpart at the two-factor level, so its domain-level and facet-level correlations were consistent. Civility (3.2) was also consistent with its counterpart at the two-factor level, showing a profile similarity of .95. Items related to trust, forgiveness, and cynicism emerged as the third factor (Trust vs. Distrust, 3.3). This factor had domain-level correlations that ranged from .41 (HEXACO Honesty-Humility) to .80 (BFI Agreeableness) and facet-level correlations that ranged from .02 (IPIP-NEO Modesty) to .84 (FI-FFM Trust), with a median of .51. The highest of these correlations were with Trust (IPIP-NEO, FI-FFM), Forgiveness and Gentleness (HEXACO), and Straightforwardness (FI-FFM). Interestingly, the correlations with the two aspects of the BFAS were similar and neither was particularly high, suggesting that the Trust factor does not fit well under either of the aspects.

At the fourth level of the analyses, Compassion (4.1) and Trust (4.3) remained consistent with Factors 3.1

TABLE 2 Example item content

Scale	Item	1.1	2.1	2.2	3.1	3.2	3.3	4.1	4.2	4.3	4.4	5.1	5.2	5.3	5.4	5.5
IPIP-NEO	I am concerned about others.	.66*	.76*	.44	.78*	.44	.34	.79*	.33	.35	.44	.78*	.30	.34	.37	.42
BFAS	I sympathize with others' feelings.	.63	.76*	.41	.77*	.40	.33	.77*	.25	.36	.45	.77*	.22	.33	.39	.40
BFAS	I can't be bothered with others' needs. (R)	.70*	.74*	.54	.73*	.53	.41	.74*	.45	.42	.48	.75*	.42	.34	.49	.37
FI-FFM	I try to take others' feelings into account.	.65*	.75*	.44	.76*	.44	.34	.75*	.24	.38	.52	.75*	.20	.34	.45	.46
BFAS	I like to do things for others.	.60	.72*	.38	.74	.37	.32	.74*	.25	.34	.41	.73*	.22	.35	.32	.40
FI-FFM	If I knew I wouldn't get caught, I'd be willing to break the law. (R)	.48	.31	.56	.26	.55	.39	.29	.65*	.34	.33	.32	.65*	.27	.39	.23
IPIP-NEO	I use flattery to get ahead. (R)	.41	.18	.56	.14	.58	.27	.16	.63*	.22	.39	.17	.63*	.19	.34	.38
HEXACO	If I want something from someone, I will laugh at that person's worst jokes. (R)	.38	.19	.49	.15	.51	.25	.18	.61	.20	.30	.19	.62*	.18	.25	.31
IPIP-NEO	I cheat to get ahead. (R)	.57	.41	.61*	.39	.63*	.32	.40	.63*	.30	.48	.43	.61*	.18	.50	.35
BFAS	I take advantage of others. (R)	.63	.48	.65*	.45	.67*	.37	.46	.63*	.35	.54	.48	.61*	.26	.52	.45
IPIP-NEO	I distrust people. (R)	.41	.37	.38	.27	.25	.69*	.27	.27	.67*	.17	.28	.27	.70*	.33	.06
FI-FFM	I generally take people at their word.	.43	.43	.35	.35	.23	.63*	.35	.19	.63*	.22	.35	.18	.69*	.29	.17
FI-FFM	I have a hard time taking someone's word. (R)	.42	.37	.40	.27	.28	.68*	.27	.27	.66*	.22	.28	.27	.68*	.36	.11
BFI	I see myself as someone who is generally trusting.	.42	.43	.34	.36	.22	.61	.35	.17	.61*	.22	.35	.16	.66*	.31	.16
FI-FFM	I think most people act in good faith.	.41	.42	.33	.35	.23	.59	.35	.20	.58	.20	.34	.19	.64*	.26	.16
BFAS	I insult people. (R)	.61	.45	.65*	.39	.63*	.49	.36	.50	.49	.60*	.41	.46	.34	.70*	.37
HEXACO	People think of me as someone who has a quick temper. (R)	.52	.37	.57	.31	.54	.47	.26	.34	.50	.59*	.30	.30	.35	.69*	.37
BFI	I see myself as someone who is sometimes rude to others. (R)	.56	.42	.58	.37	.55	.46	.34	.40	.48	.57	.38	.36	.32	.69*	.32
BFI	I see myself as someone who starts quarrels with others. (R)	.54	.38	.59	.35	.60	.35	.31	.44	.37	.60*	.35	.40	.20	.66*	.40
FI-FFM	Some people see me as insensitive. (R)	.65*	.59	.58	.54	.55	.51	.53	.42	.52	.54	.56	.38	.40	.64*	.33
FI-FFM	I try to be modest about my accomplishments.	.49	.42	.47	.43	.52	.14	.39	.29	.18	.61*	.38	.25	.18	.37	.68*
FI-FFM	I don't like to brag about my accomplishments.	.42	.32	.43	.33	.50	.07	.29	.29	.11	.58	.29	.26	.12	.31	.68*
FI-FFM	I'm not one to boast or brag.	.39	.29	.41	.30	.46	.09	.26	.25	.13	.55	.26	.22	.13	.31	.63*
FI-FFM	It is better to be modest and humble than to be proud and boastful.	.49	.43	.45	.44	.50	.14	.41	.31	.18	.56	.41	.27	.17	.35	.62*
FI-FFM	I like to show off whenever I can. (R)	.48	.31	.56	.31	.62*	.15	.30	.53	.15	.56	.30	.51	.11	.38	.59*

Note. IPIP-NEO = International Personality Item Pool-NEO; BFAS = Big Five Aspects Scale; FI-FFM = Faceted Inventory of the Five-Factor Model; BFI = Big Five Inventory. For each factor, the five highest-loading items included in the table are in boldface. An asterisk indicates that the item is among the five highest-loading items for that factor. Items not in the public domain were reprinted with permission from the copyright holders.

(Compassion) and 3.3 (Trust), as both had profile similarities of .99 with their counterparts. It was the item content related to moral behavior and interpersonal hostility (Civility, 3.2) that split to yield the second (Morality vs. Immorality, 4.2)

and fourth (Amiability vs. Rudeness, 4.4) factors. Morality (4.2) was composed of items related to immoral (i.e., dishonest and manipulative) behavior. Its domain-level correlations ranged from .38 (HEXACO Agreeableness) to .82

TABLE 3 Factor score correlations with agreeableness scales

	F1.1	F2.1	F2.2	F3.1	F3.2	F3.3	F4.1	F4.2	F4.3	F4.4	F5.1	F5.2	F5.3	F5.4	F5.5
BFI-A	.87	.81	.78	.73 ^a	.70 ^a	.80 ^b	.70 ^a	.50 ^b	.83 ^c	.74 ^a	.73 ^a	.45 ^b	.73 ^a	.83 ^c	.52 ^d
IPIP-NEO-A	.93	.82 ^a	.87 ^b	.77 ^a	.84 ^b	.70 ^c	.77 ^a	.74 ^{a,b}	.69 ^b	.76 ^a	.79 ^a	.70 ^b	.62 ^c	.75 ^d	.64 ^c
Trust	.47	.45	.41	.35 ^a	.26 ^b	.78 ^c	.36 ^a	.26 ^b	.76 ^c	.21 ^b	.36 ^a	.26 ^b	.82 ^c	.36 ^a	.12 ^d
Morality	.68	.43 ^a	.80 ^b	.37 ^a	.82 ^b	.45 ^c	.39 ^a	.87 ^b	.41 ^a	.60 ^c	.42 ^a	.86 ^b	.31 ^c	.59 ^d	.50 ^e
Altruism	.80	.88 ^a	.58 ^b	.88 ^a	.55 ^b	.51 ^b	.88 ^a	.41 ^b	.53 ^c	.57 ^c	.89 ^a	.38 ^b	.46 ^c	.57 ^d	.46 ^c
Cooperativeness	.71	.49 ^a	.78 ^b	.42 ^a	.76 ^b	.60 ^c	.39 ^a	.62 ^b	.61 ^b	.73 ^c	.44 ^a	.57 ^b	.43 ^a	.83 ^c	.47 ^a
Modesty	.35	.22 ^a	.42 ^b	.23 ^a	.50 ^b	.02 ^c	.22 ^a	.42 ^b	.02 ^c	.46 ^b	.22 ^a	.41 ^b	.01 ^c	.24 ^a	.55 ^d
Sympathy	.67	.81 ^a	.43 ^b	.83 ^a	.43 ^b	.34 ^c	.84 ^a	.32 ^b	.36 ^b	.45 ^c	.83 ^a	.29 ^b	.36 ^b	.36 ^b	.45 ^c
HEX-A	.68	.53 ^a	.70 ^b	.42 ^a	.60 ^b	.80 ^c	.37 ^a	.38 ^a	.84 ^b	.67 ^c	.40 ^a	.34 ^a	.77 ^b	.76 ^b	.48 ^c
Forgiveness	.43	.36	.42	.26 ^a	.30 ^a	.70 ^b	.24 ^{a,c}	.22 ^a	.70 ^b	.30 ^c	.25 ^a	.20 ^a	.75 ^b	.39 ^c	.23 ^a
Gentleness	.65	.56 ^a	.61 ^b	.49 ^a	.55 ^a	.64 ^b	.44 ^a	.30 ^b	.69 ^c	.66 ^c	.46 ^a	.25 ^b	.65 ^c	.66 ^c	.54 ^d
Flexibility	.52	.38 ^a	.56 ^b	.30 ^a	.50 ^b	.57 ^c	.27 ^a	.37 ^b	.58 ^c	.51 ^d	.30 ^a	.34 ^a	.50 ^b	.60 ^c	.34 ^a
Patience	.50	.34 ^a	.56 ^b	.26 ^a	.50 ^b	.55 ^b	.20 ^a	.28 ^b	.59 ^c	.60 ^c	.24 ^a	.24 ^a	.48 ^b	.68 ^c	.39 ^d
HEX-H	.67	.45 ^a	.76 ^b	.41 ^a	.79 ^b	.41 ^a	.43 ^a	.82 ^b	.37 ^a	.59 ^c	.43 ^{a,c}	.82 ^b	.37 ^a	.45 ^c	.64 ^d
Sincerity	.41	.22 ^a	.53 ^b	.18 ^a	.55 ^b	.27 ^c	.20 ^a	.64 ^b	.22 ^a	.35 ^c	.21 ^a	.65 ^b	.23 ^a	.26 ^a	.40 ^c
Fairness	.53	.35 ^a	.61 ^b	.30 ^a	.60 ^b	.44 ^c	.33 ^a	.70 ^b	.39 ^a	.38 ^a	.35 ^a	.70 ^b	.32 ^{a,d}	.44 ^c	.28 ^d
Greed Avoidance	.38	.24 ^a	.44 ^b	.22 ^a	.46 ^b	.22 ^a	.22 ^a	.46 ^b	.20 ^a	.36 ^c	.21 ^{a,c}	.47 ^b	.26 ^a	.19 ^c	.48 ^b
Modesty	.63	.51 ^a	.62 ^b	.52 ^a	.68 ^b	.23 ^c	.51 ^a	.55 ^a	.24 ^b	.66 ^c	.50 ^a	.53 ^a	.25 ^b	.41 ^c	.76 ^d
HEX-Altruism	.78	.81^a	.62^b	.80^a	.60^b	.48^c	.81^a	.48^b	.50^b	.59^c	.82^a	.45^b	.42^b	.58^c	.48^b
FI-FFM-A	.94	.80 ^a	.91 ^b	.73 ^a	.85 ^b	.79 ^c	.72 ^a	.74 ^a	.79 ^b	.78 ^b	.74 ^a	.70 ^{a,c}	.73 ^a	.79 ^b	.66 ^c
Empathy	.86	.94^a	.65^b	.93^a	.62^b	.55^c	.92^a	.44^b	.58^c	.66^d	.94^a	.40^b	.51^c	.65^d	.53^c
Trust	.53	.48	.49	.36 ^a	.35 ^a	.84 ^b	.37 ^a	.37 ^a	.82 ^b	.26 ^c	.37 ^a	.36 ^a	.87 ^b	.40 ^a	.16 ^c
Straightforwardness	.72	.46 ^a	.84 ^b	.38 ^a	.81 ^b	.64 ^c	.39 ^a	.84 ^b	.59 ^c	.62 ^c	.42 ^a	.82 ^b	.48 ^a	.70 ^c	.44 ^a
Modesty	.68	.53 ^a	.71 ^b	.53 ^a	.78 ^b	.24 ^c	.49 ^a	.53 ^a	.28 ^b	.85 ^c	.50 ^a	.48 ^a	.24 ^b	.58 ^c	.89 ^d
BFAS-A	.94	.90 ^a	.81 ^b	.88 ^a	.81 ^b	.55 ^c	.88 ^a	.65 ^b	.56 ^c	.80 ^d	.90 ^a	.60 ^b	.47 ^c	.74 ^d	.68 ^e
Compassion	.83	.93^a	.59^b	.94^a	.58^b	.47^c	.95^a	.45^b	.48^b	.58^c	.95^a	.41^b	.42^{b,d}	.55^c	.49^{c,d}
Politeness	.84	.66^a	.87^b	.62^a	.89^b	.52^c	.59^a	.72^b	.53^c	.86^d	.62^a	.68^b	.41^c	.79^d	.73^e

Note. BFI-A = Big Five Inventory-Agreeableness; IPIP-NEO-A = IPIP-NEO-Agreeableness; HEX-A = HEXACO-Agreeableness; HEX-H = HEXACO-Honesty-Humility; HEX-Altruism = HEXACO-Altruism; FI-FFM-A = Faceted Inventory of the Five-Factor Model-Agreeableness; BFAS-A = Big Five Aspects Scale-Agreeableness. The five largest facet-level correlations for each factor are in boldface. All correlations greater than $| r | = .074$ are significant at $p < .01$. At each factor level (e.g., F3.1, F3.2, F3.3), correlations in the same row with different superscripts are significantly different from one another at $p < .01$.

(HEXACO Honesty-Humility), whereas its facet-level correlations ranged from .22 (HEXACO-A Forgiveness) to .87 (IPIP-NEO Morality), with a median of .45. Its highest facet-level correlations were with Morality (IPIP-NEO), Sincerity and Fairness (HEXACO), Straightforwardness (FI-FFM), and Politeness (BFAS). Amiability (4.4) was composed of the items related to (at the negative pole) interpersonally antagonistic attitudes and behaviors. It had domain-level correlations between .59 (HEXACO Honesty-Humility) and .80 (BFAS Agreeableness). Amiability's facet scale correlations ranged from .21 (IPIP-NEO Trust) to .86 (BFAS Politeness), with a median of .58. Its highest facet correlations were with Cooperativeness (IPIP-NEO), Gentleness (HEXACO), Modesty (HEXACO, FI-FFM), Empathy (FI-FFM), and Politeness (BFAS).

In the five-factor solution, the first (Compassion, 5.1), second (Morality, 5.2), and third (Trust, 5.3) factors were consistent with those at the previous factor level, showing profile similarities of 1.00, 1.00, and .96, respectively. Amiability (4.4) split to yield the bulk of the items that loaded onto Affability (vs. Combativeness; 5.4) and Modesty (vs. Arrogance; 5.5). Affability (5.4) was composed of those items related to oppositional behaviors. It had domain-level correlations ranging from .45 (HEXACO Honesty-Humility) to .83 (BFI Agreeableness). Affability's (5.4) facet-level correlations ranged from .19 (HEXACO-H Greed Avoidance) to .83 (IPIP-NEO Cooperativeness), with a median correlation of .57. Its highest facet correlations were with Cooperativeness (IPIP-NEO), Gentleness and Flexibility (HEXACO), Straightforwardness (FI-FFM), and Politeness (BFAS).

TABLE 4 Profile similarity of agreeableness factors and factor score correlations

	F1.1	F2.1	F2.2	F3.1	F3.2	F3.3	F4.1	F4.2	F4.3	F4.4	F5.1	F5.2	F5.3	F5.4	F5.5
F1.1		.91	.92	.86	.89	.73	.84	.72	.74	.86	.86	.68	.65	.84	.72
F2.1	.89		.67	.99	.63	.61	.98	.43	.64	.69	.98	.39	.60	.65	.59
F2.2	.77	.39		.59	.98	.72	.57	.87	.71	.88	.61	.83	.59	.88	.72
F3.1	.86	.99	.35		.58	.47	.99	.37	.51	.65	.99	.32	.48	.56	.60
F3.2	.70	.33	.95	.34		.57	.56	.88	.56	.90	.59	.84	.44	.83	.78
F3.3	.38	.33	.30	.18	-.03		.46	.51	.99	.50	.49	.49	.94	.72	.27
F4.1	.85	.98	.34	.99	.33	.15		.41	.48	.59	1.00	.37	.46	.51	.54
F4.2	.41	.04	.78	.05	.84	-.11	.10		.44	.59	.43	1.00	.35	.60	.49
F4.3	.40	.37	.28	.23	-.04	.99	.19	-.19		.55	.51	.41	.95	.76	.32
F4.4	.77	.52	.83	.52	.85	.07	.47	.44	.13		.62	.53	.43	.87	.89
F5.1	.87	.98	.38	1.00	.36	.18	1.00	.12	.22	.50		.39	.46	.57	.54
F5.2	.34	-.02	.72	-.01	.79	-.14	.04	1.00	-.23	.36	.06		.33	.54	.44
F5.3	.24	.29	.09	.15	-.23	.96	.12	-.32	.96	-.08	.13	-.33		.55	.32
F5.4	.77	.53	.80	.45	.64	.59	.39	.28	.63	.80	.44	.20	.40		.55
F5.5	.53	.33	.60	.40	.77	-.37	.37	.46	-.32	.84	.38	.41	-.42		.35

Note. The lower diagonal depicts the profile similarity of the factors as measured by the Pearson correlation coefficient of the profiles from Table 3. The factors with the most similar profiles at each level of the analysis are in boldface. The upper diagonal identifies the Pearson correlations of the factor scores across each of the levels.

Modesty (5.5) was characterized by items related to vanity (e.g., “I like to show off whenever I can”) and humility (e.g., “I try to be modest about my accomplishments”). Its domain-level correlations ranged from .48 (HEXACO Agreeableness) to .68 (BFAS Agreeableness). Modesty’s (5.5) facet-level correlations ranged from .12 (IPIP-NEO Trust) to .89 (FI-FFM Modesty), with a median of .47. The highest facet correlations were with Modesty (IPIP-NEO, HEXACO, FI-FFM), Gentleness (HEXACO), and Politeness (BFAS). Each of the facets identified at this stage of the analysis demonstrated substantial convergent validity with the overall Agreeableness domain; the five facets identified correlated between .65 and .86 with the first unrotated Agreeableness factor (see Table 4). However, there do not appear to be any redundant factors. Correlations within the five-factor level ranged from .32 (Trust and Modesty) to .57 (Compassion and Affability), with a median association of .50. The uniqueness of each of the five identified factors is further supported by their unique correlational profiles.

3.2 | Criterion validity across the factor solutions

In order to evaluate factor divergence, each factor score was correlated with a number of criterion measures, including the

axes of the interpersonal circumplex, the remaining four domains of the BFI, aggression, and additional externalizing behaviors (see Table 5). Unsurprisingly, each Agreeableness factor score had a substantial relationship with the IPC Love scale, with relationships ranging in size from .50 (Modesty, 5.5) to .81 (Agreeableness, 1.1). The highest association at the five-factor level was .73 with Compassion (5.1). Correlations with IPC Dominance varied across factors from a null finding of -.05 (Agreeableness, 1.1) to a moderately sized correlation of -.33 (Modesty, 5.5). Divergent effects across factors were also found for the rest of the BFI domains. Agreeableness (1.1) demonstrated a moderate correlation with Neuroticism ($r = -.23$), but at lower levels of the analyses, correlations revealed a range of effects from -.06 (Compassion, 5.1) to -.37 (Trust; 5.3) at the five-factor level. Correlations with Extraversion ranged from small negative ($r = -.23$; Modesty, 5.5) to small positive effects ($r = .22$; Compassion, 5.1). Five-factor level associations with Openness ranged from a null association of -.03 (Morality, 5.2) to a moderate positive relation, .30 (Compassion, 5.1). Correlations with Conscientiousness consistently fell in the moderately positive range, with correlations ranging from .19 (Trust, 5.3; Modesty, 5.5) to .39 (Affability, 5.4).

Factors also diverged in their association with proactive and reactive aggression. At the five-factor level, correlations

TABLE 5 Factor score correlations with external criterion measures

	F1.1	F2.1	F2.2	F3.1	F3.2	F3.3	F4.1	F4.2	F4.3	F4.4	F5.1	F5.2	F5.3	F5.4	F5.5
IPC															
Love	.81	.76 ^a	.72 ^b	.71 ^a	.68 ^a	.61 ^b	.70 ^a	.55 ^b	.62 ^c	.66 ^{a,c}	.73 ^a	.52 ^b	.53 ^b	.69 ^a	.50 ^b
Dominance	-.05	.10 ^a	-.18 ^b	.10 ^a	-.24 ^b	.09 ^a	.12 ^a	-.17 ^b	.08 ^a	-.25 ^d	.12 ^a	-.17 ^b	.08 ^a	-.09 ^c	-.33 ^d
BFI															
Neuroticism	-.23	-.13 ^a	-.29 ^b	-.06 ^a	-.22 ^b	-.41 ^c	-.04 ^a	-.17 ^b	-.41 ^c	-.22 ^b	-.06 ^a	-.16 ^b	-.37 ^c	-.34 ^c	-.08 ^{a,b}
Extraversion	.08	.21 ^a	-.06 ^b	.20 ^a	-.13 ^b	.20 ^a	.21 ^a	-.09 ^b	.20 ^a	-.13 ^b	.22 ^a	-.09 ^b	.19 ^a	.03 ^c	-.23 ^d
Openness	.19	.29 ^a	.07 ^b	.31 ^a	.06 ^b	.09 ^b	.30 ^a	-.01 ^b	.12 ^c	.13 ^c	.30 ^a	-.03 ^b	.11 ^c	.12 ^c	.10 ^c
Conscientiousness	.35	.27 ^a	.38 ^b	.24 ^a	.36 ^b	.29 ^a	.23 ^a	.33 ^b	.28 ^{a,b}	.32 ^b	.26 ^{a,b}	.31 ^a	.19 ^b	.39 ^c	.19 ^b
Externalizing															
Proactive aggression	-.46	-.32 ^a	-.52 ^b	-.28 ^a	-.52 ^b	-.31 ^a	-.28 ^a	-.47 ^b	-.30 ^a	-.46 ^b	-.30 ^a	-.45 ^b	-.19 ^c	-.50 ^b	-.33 ^a
Reactive aggression	-.36	-.19 ^a	-.45 ^b	-.12 ^a	-.40 ^b	-.45 ^b	-.10 ^a	-.36 ^b	-.45 ^c	-.36 ^b	-.13 ^a	-.34 ^b	-.34 ^b	-.50 ^c	-.17 ^a
CAB-Drugs	-.05	.06 ^a	-.14 ^b	.09 ^a	-.13 ^b	-.12 ^b	.08 ^a	-.20 ^b	-.10 ^c	-.03 ^c	.07 ^a	-.21 ^b	-.07 ^{c,d}	-.10 ^c	.02 ^{a,d}
CAB-Crimes	-.23	-.11 ^a	-.31 ^b	-.07 ^a	-.30 ^b	-.24 ^b	-.08 ^a	-.35 ^b	-.21 ^c	-.19 ^c	-.10 ^a	-.34 ^b	-.16 ^a	-.26 ^c	-.11 ^a

Note. IPC = interpersonal circumplex; BFI = Big Five Inventory; CAB = Crime and Analogous Behavior scale. All correlations greater than $|r| = .074$ are significant at $p < .01$. At each factor level (i.e., F3.1, F3.2, F3.3), correlations in the same row with different superscripts are significantly different from one another at $p < .01$.

with proactive aggression ranged from $-.19$ (Trust, 5.3) to $-.50$ (Affability, 5.4), whereas correlations with reactive aggression ranged from $-.13$ (Compassion, 5.1) to $-.50$ (Affability, 5.4). Five-factor-level correlations with substance use ranged from $.07$ (Compassion, 5.1) to $-.21$ (Morality, 5.2), and correlations with externalizing behaviors ranged from $-.10$ (Compassion, 5.1) to $-.34$ (Morality, 5.2).

4 | DISCUSSION

The goal of the present analysis was to identify the empirical structure of Agreeableness by identifying levels of factor emergence from the domain using items from multiple self-report measures. This methodology allows for a description of the domain at varying degrees of specificity as well as a description of how each of the factors emerges from one level of the hierarchy to another. It should be acknowledged that the variance associated with Agreeableness could be parsed in any number of ways. The answer as to how to partition this variance optimally can be guided by pragmatic decisions as to which factor solution is most useful for the question at hand. For psychometric reasons, the most valuable approach will be, in many cases, one that balances parsimony with the goal of using narrow, unidimensional traits (Smith et al., 2009). However, as Smith and colleagues (2009) point out, there can be value in multidimensional constructs of varying bandwidths. For example, there is some evidence to suggest that a two-factor structure may be a better bandwidth for evaluation of genetic contributions (DeYoung et al., 2007; Jang, Livesley, Angleitner, Riemann, & Vernon, 2002). With this methodology, we provide a map

of the possible ways this variance is distributed across a full range of both unidimensional and multidimensional facets.

4.1 | Unfolding of the Agreeableness structure

In identifying the first unrotated factor, we assume that all relevant aspects of the Agreeableness domain are represented in the pool of items used. We sampled broadly from many of the most commonly used measures of Agreeableness and related traits (e.g., Honesty-Humility). We feel that the first factor of this analysis is a valid broad conceptualization of this domain, as demonstrated by its strong correlations with the domain scores from all seven scales.

Previous research on the intermediate structure of Agreeableness has found two aspects labeled *Compassion* and *Politeness* (DeYoung et al., 2007). With these two aspects, DeYoung and colleagues (2007) suggest the presence of a factor related to emotional affiliation (e.g., warmth, sympathy) and a factor related to respect for others and moral behavior (e.g., straightforwardness, cooperation). Our findings suggest a similar organization at the two-factor level. Compassion (2.1, vs. Callousness) was so labeled as the items are highly related to those found in DeYoung and colleagues' BFAS Compassion scale. Although it has been suggested that the NEO-PI-R has no good, specific markers for BFAS Compassion (DeYoung et al., 2007), the IPIP-NEO facets of Altruism and Sympathy served as excellent markers of Compassion (2.1) in the present analyses in that they loaded substantially more strongly on Compassion than the other factor that emerged at this level, Civility. In fact, Altruism, Sympathy, and Empathy appear to be defining

characteristics of the factor. It is noteworthy that Compassion stays intact in its original form from the time that it emerges at the two-factor level to the final factor solution (see Figure 2). This suggests that concern for others and sympathy for those in need is a well-defined and specific trait that is central to the Agreeableness domain. The second factor that emerged at the two-factor level was labeled *Civility* (vs. Incivility; 2.2) to reflect that the factor not only contrasts interpersonal politeness versus confrontational behavior, but also moral versus immoral or manipulative behavior, and humility versus conceitedness. It is noteworthy that all new factors beyond the third factor that emerge at each successive level of the analyses do so from the Civility factor, suggesting that its content is quite heterogeneous.

At the three-factor level, a unique trust-related factor (Trust vs. Distrust, 3.3) emerged, whereas the Compassion and Civility factors from the second level remained largely intact. The result is a three-factor structure largely consistent with that of the BFI-2 (Soto & John, 2017). Trust is defined by a willingness to trust others, a belief that others have reasonably good and sincere motives for their behavior, and a willingness to forgive others when that trust has been broken. Like Compassion, Trust was relatively stable from one solution level to the next once it emerged. Trust-related items represent a unique factor, as they have only modest associations with Compassion and Civility. This is consistent with the findings of DeYoung and colleagues (2007), who found that (NEO-PI-R) Trust had loadings of only .42 on both (BFAS) Compassion and (BFAS) Politeness. As such, the trust factor seems to exist largely outside of DeYoung and colleagues' aspects. Trust-related content differs also in that it may be an admixture of content representing both disagreeable and neuroticism-related content, as was the case here.

In the fourth level of the analyses, the heterogeneous Civility factor divided further to yield factors that we labeled *Morality* (vs. Immorality; 4.2) and *Amiability* (vs. Rudeness; 4.4). Morality is characterized by honest and moral interpersonal behaviors versus interpersonal deceit or manipulation for personal gain. It is the third stable factor to emerge in the analysis, and it is strongly related to aspects of the HEXACO Honesty-Humility domain, specifically the Sincerity and Fairness facets (Ashton et al., 2004). Amiability is defined by a combination of interpersonal traits that includes modesty, cordiality, and conflict. A four-factor structure of Agreeableness has previously been proposed as part of the FI-FFM (E. E. Simms, 2009), which includes facets of Empathy, Trust, Straightforwardness, and Modesty. The FI-FFM is generally consistent with the four-factor structure of the present model, with each of the four FI-FFM facets being representative of one of the four identified facets. However, examination of item content suggests that the FI-FFM Modesty facet may be a somewhat narrowly defined trait given

that Amiability contains items related to modesty as well as a confrontational (vs. conflict-avoidant) interpersonal style.

The fifth factor solution is the first in which all of the factors appear to contain reasonably homogeneous item content. At this level, the mixed item content of Amiability yields two factors that we labeled *Modesty* (vs. Arrogance; 5.5) and *Affability* (vs. Combativeness; 5.4). Modesty is reasonably straightforward, as the three modesty facet scales included in these analyses (i.e., IPIP-NEO Modesty, HEXACO Modesty, FI-FFM Modesty) are the primary correlates of the factor. Affability is a coherent factor that contains all items related to interpersonal cooperation (vs. confrontation). The structure at this stage of the analysis is quite similar to the six-facet model of the FFM, as operationalized by the NEO-PI-R and IPIP-NEO-PI-R, as the IPIP-NEO facets of Trust, Morality, Cooperation, and Modesty are all represented by highly related factors at this point. However, the present analyses suggest that a sixth facet is unnecessary, as IPIP-NEO Altruism and Sympathy facets are better represented by a single factor. Along with being unidimensional in content, the five facets at this stage of analysis are nonredundant, and they show some degree of discriminant validity in relation to the rest of the B5 domains and externalizing behaviors.

When interpreting these results, it is important to consider the methodology's dependence on the initial pool of items used. Like all factor solutions, the validity of our proposed structure is contingent on the inclusion of items that effectively cover the Agreeableness domain. If there are substantial components of the domain that are not well represented in the pool of items, those components would be unable to emerge as a unique factor. We relied on commonly used measures of Agreeableness (rather than the English lexicon) for our initial item pool, which has a number of implications for the identified factors. The approach allows us to evaluate the structure of Agreeableness as it is generally conceptualized in the literature, but the relative preponderance of certain items (e.g., those related to altruism), which are present in nearly all of the scales, makes them highly predictive of the single Agreeableness factor, whereas the relative underrepresentation of other factors (e.g., HEXACO Greed Avoidance) results in smaller loadings on the domain. The item selection method is particularly important in relation to the HEXACO, which accounts for Agreeableness-related variance in a relatively unique manner (see Ashton et al., 2014, for a relevant review), and some measures of Agreeableness do not effectively capture the Honesty-Humility domain (Ashton & Lee, 2005; Miller, Gaughan, Maples, & Price, 2011). However, Honesty-Humility content is present in IPIP-NEO scales (Morality, Modesty), FI-FFM scales (Straightforwardness, Modesty), and of course the HEXACO items themselves. Together, Honesty-Humility-related scales represented nearly 33% of the total item pool (i.e., 34 out of 104 items). It therefore seems unlikely that the lack of a clear

Honesty-Humility emergence could be attributable to underrepresentation in the item pool. We do acknowledge that some potentially appropriate items were excluded, including IPIP-NEO-PI-R items related to Warmth (from the domain of Extraversion), IPIP-NEO-PI-R items related to Angry Hostility (from the domain of Neuroticism), and BFAS Volatility items. It is possible that these exclusions had some effect on the centrality of the HEXACO's conceptualization of Agreeableness, which incorporates anger. These measures of interpersonal warmth and irritability could have reasonably been included in these analyses along with clinically oriented assessments (e.g., the Personality Inventory for DSM-5 Antagonism domain; Krueger, Derringer, Markon, Watson, & Skodol, 2012). Such exclusions were necessary due to practical limitations on survey length, and the selection of only scales whose primary loadings are on the Agreeableness domain seems a reasonable demarcation consistent with the facet-level examinations that have been conducted for Extraversion and Conscientiousness (Roberts et al., 2005; Watson et al., 2015). That said, if additional measures were to be included in the item pool, the identified factor structure could look, or at least unfold, in a different manner.

4.2 | Criterion validity

Agreeableness factor scores were correlated with a range of relevant criterion variables. Of the 10 external associations assessed, five include divergent correlations or a combination of significant and null associations across facets. For instance, the first unrotated Agreeableness factor manifested a null association with Dominance of the IAS (Wiggins, 1995). However, when the Agreeableness domain is parsed into facets, some divergent associations emerge. Compassion, Trust, Morality, and Affability all have modest associations with Dominance, whereas Modesty has a moderately sized negative association. This is consistent with the notion that immodesty/grandiosity may be a mixture of low Agreeableness and high Extraversion/Dominance (e.g., Crowe, Carter, Campbell, & Miller, 2016; Paulhus, 2001). Extraversion's association with Agreeableness shows a similar pattern, with facets of Compassion and Modesty demonstrating correlations of similar size but opposite valences. Agreeableness is related to aggression (e.g., Bettencourt, Talley, Benjamin, & Valentine, 2006; Jones et al., 2011), but significant variability in the size of the associations at the facet level could mask particular variables of interest if the domain alone was assessed. In the present analyses, facet-level correlations ranged from $-.19$ to $-.50$ and $-.13$ to $-.50$ for proactive and reactive aggression, respectively. These data suggest that Affability is a particularly relevant facet for both forms of aggression, whereas Compassion is more specifically relevant to proactive aggression and Trust is more

specifically relevant to reactive aggression. Only through fine-grained personality distinctions such as these can we begin to identify the specific cognitions and motivations associated with such outcomes.

4.3 | Assessment implications

Despite the different processes through which current facet-level measures of Agreeableness were developed, it seems that many do a reasonable job of describing certain levels of the domain. Nearly all extant Agreeableness scales find counterparts within our factor solutions. The BFAS seems to be largely congruent with the two-factor level, the three-factor level is generally consistent with the structure of the BFI-2, the FI-FFM facets of Agreeableness each have primary loadings on one factor at the four-facet level, and although Sympathy and Altruism do not split to become unique facets, the IPIP-NEO-PI-R-based representation includes primary indicators of each of the five facets identified in the final level of our analyses. Even the HEXACO domains of Agreeableness and Honesty-Humility can be found across different solutions, although the two-factor level of our analyses does not suggest an Honesty-Humility and Agreeableness split (both domains have stronger loadings on the Civility factor). Certain HEXACO facets did demonstrate relatively weaker associations with the individual factors identified here, however. For example, Greed Avoidance and Flexibility, facets of the HEXACO Honesty-Humility and Agreeableness scales, respectively, are the only facet-level measures used in these analyses that failed to fall among the top five strongest indicators for any factor. However, the HEXACO's interstitial Altruism scale is among the primary indicators for our identified Compassion factor throughout each of the factor solutions.

5 | LIMITATIONS AND FUTURE DIRECTIONS

Although this study has a number of strengths—including a large sample size, a diverse range of Agreeableness measures, and a focus on the item-level, rather than scale-level, structure—a number of limitations should be acknowledged.

First, although samples collected from MTurk tend to be more ethnically diverse than a typical undergraduate research pool population (Buhrmester, Kwang, & Gosling, 2011; Miller, Crowe, Weiss, Maples-Keller, & Lynam, 2017), participants were relatively homogeneous and limited to individuals residing in the United States. As a result of this White American-centric population, we cannot presume that the identified structure will replicate in other countries and cultural contexts. As with previous research conducted on Extraversion and Conscientiousness, additional

comprehensive analyses will need to be conducted in order to test the extent to which the facet-level structure generalizes. Notably, there is strong evidence that these personality dimensions are largely culture nonspecific (Ashton & Lee, 2001; Church & Katigbak, 2005; McCrae, 2001), but the extent to which the more specific facets are cross-culturally consistent is unclear.

The sole reliance on self-report measures is an additional limitation, as it is possible that the strength of our identified associations between facets and criterion measures is inflated due to shared method variance, although these effects are offset by the less than perfect reliability of these items (Chan, 2009). Extreme response styles may have inflated correlations, but analyses conducted while controlling for response bias indicated that there was no meaningful effect on the identified factor structure. Our reliance on an online data collection approach is another potential limitation. Although there are a number of advantages to the use of crowdsourcing approaches like MTurk (see Chandler & Shapiro, 2016; Miller et al., 2017, for reviews), the equivalency of these assessments across in-person administrations and online administrations has not been formally assessed and such equivalency cannot necessarily be assumed (see Bagby, Ayearst, Morariu, Watters, & Taylor, 2014, for a discussion of these issues). Nonetheless, previous examinations have suggested that personality measures tend to show equivalence across these administration modalities (Meade, Michels, & Lautenschlager, 2007). Future research could also benefit from the collection of additional criterion variables. The present research chose to emphasize a comprehensive set of Agreeableness items over wider diversity in criterion variables. While the utility of the domains could be initially assessed with measures of the interpersonal circumplex, psychopathy, and externalizing behaviors, future work would benefit from examining the facets in relation to a wider array of relevant constructs.

6 | CONCLUSIONS

The present study demonstrates how Agreeableness unfolds at various levels of factor specificity and the value of using an empirically derived lower-order structure in evaluating important Agreeableness-related correlates (e.g., aggression). With the value of the B5/FFM domains well established, it is important for research to move beyond examinations of domain-level associations only, which are necessarily limited in their ability to identify specific associations with outcome or criterion variables. Although there is certainly important utility and value across the range of measurement specificity, we believe that many occasions call for increasingly specific personality measurements that allow for more precise personality-based models. We believe the current analyses

and model of Agreeableness represent an important step in the progression toward a comprehensive, flexible, and empirically informed conceptualization of this domain.

ACKNOWLEDGMENT

The author(s) received no financial support for the research, authorship, and/or publication of this article.

CONFLICT OF INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

NOTES

¹ Note that Goldberg (2006) was not the first to utilize this approach. Saucier (2003) utilized the method, and Di Blas and Forzi (1998) utilized a similar approach. We refer readers to Goldberg's (2006) article not only because it is the reference in which the "Bass-Ackwards" term was coined, but also because it provides a thorough explanation of the procedure and its potential value.

² In order to evaluate the generalizability of our findings across sexes, measurement invariance was evaluated (Vandenberg & Lance, 2000). Analyses support configural and metric invariance across sex, indicating that factor structure and loadings are equivalent for males and females. See the online appendix for complete results of these analyses.

³ The descriptive statistics for each of the Agreeableness measures used are included in Table 1.

⁴ Due to extreme redundancy, some items were removed prior to the factor analyses. A complete description of this process is provided in the Data Analysis section.

⁵ Eighteen items were removed because some items were present in multiple pairs of items. When selecting items to remove from redundant pairs, the item that had the least total redundancies was retained. For example, if Item 1 and Item 2 were redundant, but Item 1 had only one redundancy while Item 2 was also redundant with Item 3, Item 1 would be retained. In this way, two redundancies would be removed by dropping only one item. If items had the same number of redundancies, items were chosen randomly. By prioritizing retention of the largest number of items possible, we hoped to maximize our coverage of the Agreeableness domain. The .65 cut-off for overlapping items was rationally selected post hoc. We feel this is a reasonable cut-off that both maximizes item content and minimizes the risk of bloated specific factors. Removal of more items (by using a cut-off less than .65) would risk removing items with unique contributions to factor structure. Removal of fewer items (using a cut-off greater than .65) would be unlikely to affect results (see note 8).

⁶ A list of the final pool of items is available from the first author upon request.

⁷ Because acquiescence (inflated correlations among similarly keyed items) is a concern in item-level factor analyses, we conducted additional analyses. While all factors in the five-factor solution contained both positively and negatively keyed items, some factors did have a relative preponderance of negatively keyed items (i.e., Factor 5.2: 15

of 18 items; Factor 5.4: 22 of 25 items). In order to verify that our identified factor structure represents true item associations, analyses to evaluate the effect of acquiescence were conducted. A series of 16 opposite-item pairs from within the BFI has been previously identified (Soto, John, Gosling, & Potter, 2008). Each participant's level of acquiescence was indexed by his or her mean response to these 16 items. Simple regressions were then run predicting each of the 104 Agreeableness items with the residuals saved. Factor analyses were conducted on the saved residuals from these regressions. Tucker's congruence coefficients were calculated for each factor at all five levels of the analyses (Lorenzo-Seva & Ten Berge, 2006). Factor congruence values ranged from .98 to greater than .99 (median = .996). This indicates acquiescence had no effect on factor structure.

⁸ All solutions were also evaluated using all 131 Agreeableness items. The final five-factor solution from the full pool of items was consistent with that reported here. The three-factor and four-factor solutions were also generally consistent with our current findings. One substantial difference that did arise was at the two-factor level. When all 131 items were used, Factor 2.1 emerged as a heterogeneous mixture of what we identify as Compassion and Civility, whereas Factor 2.2 emerged as a Trust (vs. Dis-trust) factor that was generally consistent with our reported Factor 3.3.

⁹ We adopted a labeling system in which factors whose primary items remained consistent from one level to another also have consistent labels across those levels.

ORCID

Michael L. Crowe  <http://orcid.org/0000-0002-8834-067X>
 Donald R. Lynam  <http://orcid.org/0000-0001-8306-498X>
 Joshua D. Miller  <http://orcid.org/0000-0003-1280-2863>

REFERENCES

- Ashton, M. C., & Lee, K. (2001). A theoretical basis for the major dimensions of personality. *European Journal of Personality*, 15, 327–353. <https://doi.org/10.1002/per.417>
- Ashton, M. C., & Lee, K. (2005). Honesty-humility, the Big Five, and the five-factor model. *Journal of Personality*, 73, 1321–1353. <https://doi.org/10.1111/j.1467-6494.2005.00351.x>
- Ashton, M. C., Lee, K., & De Vries, R. E. (2014). The HEXACO honesty-humility, agreeableness, and emotionality factors: A review of research and theory. *Personality and Social Psychology Review*, 18, 139–152. <https://doi.org/10.1177/1088868314523838>
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., De Vries, R. E., Di Blas, L., ... De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholinguistic studies in seven languages. *Journal of Personality and Social Psychology*, 86, 356–366. <https://doi.org/10.1037/0022-3514.86.2.356>
- Bagby, R. M., Ayearst, L. E., Morariu, R. A., Watters, C., & Taylor, G. J. (2014). The Internet administration version of the 20-item Toronto Alexithymia Scale. *Psychological Assessment*, 26, 16–22. <https://doi.org/10.1037/a0034316>
- Bettencourt, B., Talley, A., Benjamin, A. J., & Valentine, J. (2006). Personality and aggressive behavior under provoking and neutral conditions: A meta-analytic review. *Psychological Bulletin*, 132, 751–777. <https://doi.org/10.1037/0033-2909.132.5.751>
- Block, J. (1995). Going beyond the five factors given: Rejoinder to Costa and McCrae (1995) and Goldberg and Saucier (1995). *Psychological Bulletin*, 117, 226–229. <https://doi.org/10.1037/0033-2909.117.2.226>
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3–5. <https://doi.org/10.1177/1745691610393980>
- Carlo, G., Okun, M. A., Knight, G. P., & De Guzman, M. R. T. (2005). The interplay of traits and motives on volunteering: Agreeableness, extraversion and prosocial value motivation. *Personality and Individual Differences*, 38, 1293–1305. <https://doi.org/10.1016/j.paid.2004.08.012>
- Chan, D. (2009). So why ask me? Are self-report data really that bad? In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 309–336). New York, NY: Routledge/Taylor & Francis Group. https://doi.org/10.1111/j.1744-6570.2009.01160_5.x
- Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, 12, 53–81. <https://doi.org/10.1146/annurev-clinpsy-021815-093623>
- Church, A. T., & Katigbak, M. S. (2005). Personality structure across cultures: Indigenous and cross-cultural perspectives. In A. Elias, S. E. Hampson, & B. De Raad (Eds.), *Advances in personality psychology* (Vol. 2, pp. 124–157). New York, NY: Psychology Press. <https://doi.org/10.4324/9780203000953>
- Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, 4, 5–13. <https://doi.org/10.1037/1040-3590.4.1.5>
- Costa, P. T., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences*, 12, 887–898. [https://doi.org/10.1016/0191-8869\(91\)90177-D](https://doi.org/10.1016/0191-8869(91)90177-D)
- Crowe, M. L., Carter, N. T., Campbell, W. K., & Miller, J. D. (2016). Validation of the Narcissistic Grandiosity Scale and creation of reduced item variants. *Psychological Assessment*, 28, 1550–1560. <https://doi.org/10.1037/pas0000281>
- DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin*, 124, 197–229. <https://doi.org/10.1037/0033-2909.124.2.197>
- 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, 880–896. <https://doi.org/10.1037/0022-3514.93.5.880>
- Di Blas, L., & Forzi, M. (1998). An alternative taxonomic study of personality-descriptive adjectives in the Italian language. *European Journal of Personality*, 12, 75–101. [https://doi.org/10.1002/\(SICI\)1099-0984](https://doi.org/10.1002/(SICI)1099-0984)
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417–440. <https://doi.org/10.1146/annurev.ps.41.020190.002221>
- Gleason, K. A., Jensen-Campbell, L. A., & South Richardson, D. (2004). Agreeableness as a predictor of aggression in adolescence. *Aggressive Behavior*, 30, 43–61. <https://doi.org/10.1002/ab.20002>

- Goldberg, L. R. (2006). Doing it all bass-ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality*, 40, 347–358. <https://doi.org/10.1016/j.jrp.2006.01.001>
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40, 84–96. <https://doi.org/10.1016/j.jrp.2005.08.007>
- Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person \times situation perspective. *Journal of Personality and Social Psychology*, 93, 583–599. <https://doi.org/10.1037/0022-3514.93.4.583>
- Graziano, W. G., & Tobin, R. M. (2017). Agreeableness and the five-factor model. In T. A. Widiger (Ed.), *The Oxford handbook of the five-factor model*. <https://doi.org/10.1093/oxfordhb/9780199352487.013.17>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185. <https://doi.org/10.1007/BF02289447>
- 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, 83–101. [https://doi.org/10.1016/S0031-8869\(01\)00137-4](https://doi.org/10.1016/S0031-8869(01)00137-4)
- Jensen-Campbell, L. A., Adams, R., Perry, D. G., Workman, K. A., Furrella, J. Q., & Egan, S. K. (2002). Agreeableness, extraversion, and peer relations in early adolescence: Winning friends and deflecting aggression. *Journal of Research in Personality*, 36, 224–251. <https://doi.org/10.1006/jrpe.2002.2348>
- Jensen-Campbell, L. A., & Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal conflict. *Journal of Personality*, 69, 323–362. <https://doi.org/10.1111/1467-6494.00148>
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research. <https://doi.org/10.1037/t07550-000>
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). New York, NY: Guilford Press.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York, NY: Guilford Press.
- Jones, S. E., Miller, J. D., & Lynam, D. R. (2011). Personality, antisocial behavior, and aggression: A meta-analytic review. *Journal of Criminal Justice*, 39, 329–337. <https://doi.org/10.1016/j.jcrimjus.2011.03.004>
- Kenny, D. A. (2015). Measuring model fit. Retrieved from <http://davidakenny.net/cm/fit.htm>
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine*, 42, 1879–90. <https://doi.org/10.1017/S0033291711002674>
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, 39, 329–358. https://doi.org/10.1207/s15327906mbr3902_8
- Lee, K., & Ashton, M. C. (2006). Further assessment of the HEXACO Personality Inventory: Two new facet scales and an observer report form. *Psychological Assessment*, 18, 182–191. <https://doi.org/10.1037/1040-3590.18.2.182>
- Lorenzo-Seva, U., & Ten Berge, J. M. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, 2, 57–64. <https://doi.org/10.1027/1614-2241.2.2.57>
- Lynam, D. R., Gaughan, E. T., Miller, J. D., Miller, D. J., Mullins-Sweatt, S., & Widiger, T. A. (2011). Assessing the basic traits associated with psychopathy: Development and validation of the Elemental Psychopathy Assessment. *Psychological Assessment*, 23, 108–124. <https://doi.org/10.1037/a0021146>
- Lynam, D. R., & Widiger, T. A. (2001). Using the five-factor model to represent the DSM-IV personality disorders: An expert consensus approach. *Journal of Abnormal Psychology*, 110, 401–412. <https://doi.org/10.1037/0021-843X.110.3.401>
- Maples, J. L., Guan, L., Carter, N. T., & Miller, J. D. (2014). A test of the International Personality Item Pool representation of the Revised NEO Personality Inventory and development of a 120-item IPIP-based measure of the five-factor model. *Psychological Assessment*, 26, 1070–1084. <https://doi.org/10.1037/pas0000004>
- Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, 18, 257–284. <https://doi.org/10.1037/a0032773>
- McCrae, R. R. (2001). Trait psychology and culture: Exploring intercultural comparisons. *Journal of Personality*, 69, 819–846. <https://doi.org/10.1111/1467-6494.696166>
- Meade, A. W., Michels, L. C., & Lautenschlager, G. J. (2007). Are Internet and paper-and-pencil personality tests truly comparable? An experimental design measurement invariance study. *Organizational Research Methods*, 10, 322–345. <https://doi.org/10.1177/1094428106289393>
- Miller, J. D., Crowe, M. L., Weiss, B., Maples-Keller, J. L., & Lynam, D. R. (2017). Using online, crowdsourcing platforms for data collection in personality disorder research: The example of Amazon's Mechanical Turk. *Personality Disorders: Theory, Research, and Treatment*, 8, 26–34. <https://doi.org/10.1037/per0000191>
- Miller, J. D., Gaughan, E. T., Maples, J., & Price, J. (2011). A comparison of agreeableness scores from the Big Five Inventory and the NEO PI-R: Consequences for the study of narcissism and psychopathy. *Assessment*, 18, 335–339. <https://doi.org/10.1177/1073191111411671>
- Miller, J. D., & Lynam, D. (2001). Structural models of personality and their relation to antisocial behavior: A meta-analytic review. *Criminology*, 39, 765–798. <https://doi.org/10.1111/j.1745-9125.2001.tb00940.x>
- Miller, J. D., & Lynam, D. R. (2003). Psychopathy and the five-factor model of personality: A replication and extension. *Journal of Personality Assessment*, 81, 168–178. https://doi.org/10.1207/S15327752JPA8102_08
- Miller, J. D., Lyman, D. R., Widiger, T. A., & Leukefeld, C. (2001). Personality disorders as extreme variants of common personality dimensions: Can the five-factor model adequately represent psychopathy? *Journal of Personality*, 69, 253–276. <https://doi.org/10.1111/1467-6494.00144>
- Naragon-Gainey, K., Watson, D., & Markon, K. E. (2009). Differential relations of depression and social anxiety symptoms to the

- facets of extraversion/positive emotionality. *Journal of Abnormal Psychology*, 118, 299–310. <https://doi.org/10.1037/a0015637>
- Osborne, J. W., Costello, A. B., & Kellow, J. T. (2008). Best practices in exploratory factor analysis. In J. W. Osborne (Ed.), *Best Practices in Quantitative Methods* (pp. 86–99). Thousand Oaks, CA: SAGE Publications Ltd. Retrieved from <https://doi.org/10.4135/9781412995627>
- Paulhus, D. L. (2001). Normal narcissism: Two minimalist accounts. *Psychological Inquiry*, 12, 228–230. https://doi.org/10.1207/S15327965PLI1204_2
- Perugini, M., & Gallucci, M. (1997). A hierarchical faceted model of the Big Five. *European Journal of Personality*, 11, 279–301. [https://doi.org/10.1002/\(SICI\)1099-0984\(199711\)11:4<279::AID-PER282>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1099-0984(199711)11:4<279::AID-PER282>3.0.CO;2-F)
- Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., ... Liu, J. (2006). The Reactive-Proactive Aggression Questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior*, 32, 159–171. <https://doi.org/10.1002/ab.20115>
- Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, 58, 103–139. <https://doi.org/10.1111/j.1744-6570.2005.00301.x>
- Samuel, D. B., & Widiger, T. A. (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: A facet level analysis. *Clinical Psychology Review*, 28, 1326–1342. <https://doi.org/10.1016/j.cpr.2008.07.002>
- Saucier, G. (2003). Factor structure of English-language personality type-nouns. *Journal of Personality and Social Psychology*, 85, 695–708. <https://doi.org/10.1037/0022-3514.85.4.695>
- Saucier, G. (2009). Recurrent personality dimensions in inclusive lexical studies: Indications for a Big Six structure. *Journal of Personality*, 77, 1577–1614. <https://doi.org/10.1111/j.1467-6494.2009.00593.x>
- Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology*, 76, 613–627. <https://doi.org/10.1037/0022-3514.76.4.613>
- Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical review. *Personality and Social Psychology Review*, 12, 248–279. <https://doi.org/10.1177/1088868308319226>
- Simms, E. E. (2009). *Assessment of the facets of the five factor model: Further development and validation of a new personality measure* (Doctoral dissertation, University of Iowa). Retrieved from <http://search.proquest.com/docview/304901508>
- Simms, L. J., & Watson, D. (2007). The construct validation approach to personality scale construction. In R. W. Robins, R. C. Fraley, & R. F. Kruger (Eds.), *Handbook of research methods in personality psychology* (pp. 240–258). New York, NY: Guilford Press.
- Smith, G. T., McCarthy, D. M., & Zapolski, T. C. B. (2009). On the value of homogeneous constructs for construct validation, theory testing, and the description of psychopathology. *Psychological Assessment*, 21, 272–284. <https://doi.org/10.1037/a0016699>
- Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality*, 43, 84–90. <https://doi.org/10.1016/j.jrp.2008.10.002>
- Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113, 117–143. <https://doi.org/10.1037/pspp0000096>
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, 94, 718–737. <https://doi.org/10.1037/0022-3514.94.4.718>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70. <https://doi.org/10.1177/109442810031002>
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321–327. <https://doi.org/10.1007/BF02293557>
- Watson, D., Stasik, S. M., Ellickson-Larew, S., & Stanton, K. (2015). Extraversion and psychopathology: A facet-level analysis. *Journal of Abnormal Psychology*, 124, 432–446. <https://doi.org/10.1037/abn0000051>
- Watson, D., Stasik, S. M., Ro, E., & Clark, L. A. (2013). Integrating normal and pathological personality: Relating the DSM-5 trait-dimensional model to general traits of personality. *Assessment*, 20, 312–326. <https://doi.org/10.1177/1073191113485810>
- Wiggins, J. S. (1995). *Interpersonal Adjective Scales: Professional manual*. Odessa, FL: Psychological Assessment Resources. <https://dx.doi.org/10.1037/02267-000>
- Witt, L. A., Burke, L. A., Barrick, M. A., & Mount, M. K. (2002). The interactive effects of conscientiousness and agreeableness on job performance. *Journal of Applied Psychology*, 87, 164–169. <https://doi.org/10.1037/0021-9010.87.1.164>

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Appendix

Table A1

Supplemental Table 1

Supplemental Table 2

Supplemental Table 3

Supplemental Table 4

Supplemental Figure 1

How to cite this article: Crowe ML, Lynam DR, Miller JD. Uncovering the structure of agreeableness from self-report measures. *Journal of Personality*. 2018;86:771–787. <https://doi.org/10.1111/jopy.12358>