


Extracting Agency and Communion From the Big Five: A Four-Way Competition

Theresa M. Entringer^{1,2} , Jochen E. Gebauer^{2,3},
and Delroy L. Paulhus⁴

Assessment
2022, Vol. 29(6) 1216–1235
© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10731911211003978
journals.sagepub.com/home/asm



Abstract

Agency and communion are the two fundamental content dimensions in psychology. The two dimensions figure prominently in many psychological realms (personality, social, self, motivational, cross-cultural, etc.). In contemporary research, however, personality is most commonly measured within the Big Five framework. We developed novel agency and communion scales based on the items from the most popular nonpropriety measure of the Big Five—the Big Five Inventory. We compared four alternative scale-construction methods: *expert rating*, *target scale*, *ant colony*, and *brute force*. Across three samples ($N_{\text{total}} = 942$), all methods yielded reliable and valid agency and communion scales. Our research provides two main contributions. For psychometric theory, it extends knowledge on the four scale-construction methods and their relative convergence. For psychometric practice, it enables researchers to examine agency and communion hypotheses with extant Big Five Inventory data sets, including those collected in their own labs as well as openly accessible, large-scale data sets.

Keywords

agency, communion, Big Five, scale construction

Two fundamental dimensions have loomed large in psychology (Abele & Wojciszke, 2007). In the terminology of Hogan (1982), one dimension—agency—refers to “getting ahead”; the other—communion—refers to “getting along.” Under various labels, these two dimensions have emerged in personality psychology (dominance vs. nurturance), social psychology (competence vs. warmth), psychology of self (independent vs. interdependent selves), motivation science (agentic vs. communal values), gender studies (instrumental vs. expressive styles), cross-cultural psychology (individualism vs. collectivism), and many more (reviewed in Abele & Wojciszke, 2018; Wiggins, 1991). Thus, the two dimensions are a central point of contact between otherwise segregated realms of psychology (Judd et al., 2005; Martin & Slepian, 2020). As such, the two dimensions also help majorly to integrate those segregated realms of psychology into one unified psychological science—a timely endeavor (Brewer, 2013; Fournier et al., 2015).¹

Despite the indisputable significance of the two fundamental dimensions, personality psychologists are far more likely to measure personality within the Big Five framework. Although several popular Big Five measures are available (e.g., Costa & McCrae, 2008; Goldberg, 1999; Gosling et al., 2003; Saucier, 1994; Woods & Hampson, 2005), the *Big Five Inventory* (BFI; John & Srivastava, 1999) is the most popular nonproprietary one. The present

research sought to extract the two fundamental dimensions from the BFI with four scale construction methods: The *expert rating (ER) method* (Hase & Goldberg, 1967; Lynam & Widiger, 2001), the *target scale (TS) method* (Goldberg, 1999; Johnson, 2014), the *ant colony (AC) method* (Dorigo & Gambardella, 1997; Leite et al., 2008), and the *brute force (BF) method* (Russell et al., 2004; Stanton et al., 2002).

We expected our efforts to succeed for theoretical and empirical reasons. Theoretically, the Big Five and the two fundamental dimensions are both exhaustive taxonomies, albeit on different levels of abstraction: As the two fundamental dimensions, agency and communion are even more basic or broader than the Big Five domains (Paulhus & John, 1998; Wiggins, 2003). Empirically, some extant research has successfully assessed agency and communion with items from another Big Five measure—the NEO-PI-R (Traupman et al., 2009) and its most recent update (NEO-PI-3; Louie et al., 2018). We expected two main contributions from the

¹German Institute for Economic Research, Berlin, Germany

²University of Mannheim, Mannheim, Baden-Württemberg, Germany

³University of Copenhagen, Copenhagen, Denmark

⁴University of British Columbia, Vancouver, British Columbia, Canada

Corresponding Author:

Theresa M. Entringer, German Institute for Economic Research (DIW Berlin), Mohrenstrasse 58, Berlin D-10117, Germany.
Email: theresa.entringer@gmail.com

present research. First, a comparative evaluation of the four scale-construction methods listed above. Such an evaluation is heretofore unprecedented, but it is in the spirit of Hase and Goldberg (1967), who pioneered the comparison of (earlier) scale-construction methods. Second, the creation of new agency and communion scales based on BFI items. Those “BFI-Agency-Communion Scales” will allow researchers to examine agency and communion hypotheses with extant BFI data sets, including data sets previously collected in their own labs and openly accessible, large-scale archival data sets.

The remainder of this Introduction is structured as follows. First, we elaborate on the two fundamental dimensions of agency and communion, including some of their most defining covariates. Second, we describe our four complementary scale construction methods. Finally, we provide an overview of the present empirical work.

The Two Fundamental Dimensions of Agency and Communion

Agency and communion are the two most basic or broadest content dimensions in psychology (Abele & Wojciszke, 2018; Wiggins, 2003). The fact that different researchers have used different terms for those dimensions may have masked their omnipresence. Here, we opt for Bakan’s (1966) terms, *agency* and *communion*, because they have had the greatest historical influence. The basic trait of agency manifests itself in a broad constellation of more specific traits, such as (elements of) dominance, competence, and extraversion. Likewise, the basic trait of communion manifests itself in a broad constellation of more specific traits, such as (elements of) morality, warmth, and agreeableness (Gebauer et al., 2013; Paulhus & John, 1998; Wiggins, 2003, see Footnote 1).

To elucidate the nature of agency and communion, scholars have often referred to their cardinal features. Bakan (1966), for instance, stated that “Agency manifests itself in the formation of separations; communion in the lack of separations” (p. 15). Wiggins (1991) noted that “Agency refers to the condition of being a differentiated individual. . . . Communion refers to the condition of being part of a larger social or spiritual entity” (p. 89). Abele and Wojciszke (2007) corroborated this conceptual link: “agency arises from strivings to individuate and expand the self. . . [C]ommunion arises from strivings to integrate the self in a larger social unit” (p. 751). The most recent framing compares *social assimilation* versus *social contrast*—the tendency to swim with the social tide versus against it (Gebauer et al., 2013).

An essential element of agency and communion is positive self-evaluation: High agency and high communion both carry positive valence (Paulhus, 2018). Thus, both are

correlated with high self-esteem (Abele et al., 2016; Gebauer et al., 2013). Of the two, agency is typically a stronger correlate of self-esteem than is communion (Abele et al., 2016; Wojciszke et al., 2011). Even though there are some exceptions (Bi et al., 2013; Gebauer et al., 2013).

Scale Construction Methods

We applied four alternative scale-construction methods to create our BFI-Agency-Communion Scales: ER (Hase & Goldberg, 1967; Lynam & Widiger, 2001), TS (Goldberg, 1999; Johnson, 2014), AC (Dorigo & Gambardella, 1997; Leite et al., 2008), and BF (Russell et al., 2004; Stanton et al., 2002). Next, we elaborate on those four methods, including their strengths and limitations. For the moment, we merely outline each method’s general steps; the Method section describes our unique adaptations.

Expert Rating Method. The ER method (Hase & Goldberg, 1967; Lynam & Widiger, 2001) involves three steps. First, several experts are asked to rate all candidate items (here: the 44 BFI-items) according to their saturation with content from the target construct (here: agency and communion). Second, those ERs are averaged for each candidate item (assuming sufficient interrater reliability). Finally, items that surpass some saturation threshold (here: mean expert-ratings ≥ 5 on a 7-point rating-scale) are averaged into the final composites (here: ER-Agency and ER-Communion Scales).

Strengths and limitations. The ER method has two main strengths: (a) the method is one of the simplest and most transparent techniques to create new scales, if experts on the target construct are available to judge the candidate items and if those items cover full construct breadth. (b) Applying the method ensures that the final scales cover current theoretical understanding of the target construct. The method, however, also has two main limitations: (a) the quality of the resultant scales depends on the experts’ idiosyncratic understanding of the target construct. (b) The resultant scales may suffer from low reliabilities, because the items are judged in isolation, rather than in concert with each other.

Target Scale Method. This method was initially proposed to construct a variety of scales from the large International Personality Item Pool (IPIP; Goldberg et al., 2006; Johnson, 2014).^{2,3} It involves two steps. First, all candidate items (here: the 44 BFI-items) are correlated with a TS (here: previously established agency and communion scales). Second, candidate items are selected if they possess high correlations with the TS. In short, the rationale is to maximize the item-to-target correlation.

Strength and limitations. The TS method has two main strengths: (a) the scale constructor does not need to be an expert regarding the target construct. (b) The resulting scale usually possesses very high convergent validity because, by definition, the items correlate highly with extant measures of the same construct (i.e., the TS). The TS method also has two main limitations: (a) the performance of *each candidate item* is considered in isolation from other candidate items. Candidate items that perform well (poorly) on their own, however, can perform comparatively poorly (well) when combined into a single score (Stanton, 2000). (b) The TS method uses only a single criterion for scale construction—namely, the correlation between each candidate item and the TS. Thus, the final scale entirely depends on the validity of the chosen TS (Loevinger, 1957).

Ant Colony Method. The AC method is a relatively new scale construction method based on a heuristic that mimics the foraging behavior of ants (Dorigo & Gambardella, 1997; Leite et al., 2008). The AC method involves four steps. First, it selects a predefined number of *candidate scales* (here: e.g., candidate AC-Agency Scales). The items of those candidate scales are random subsets of the candidate items (here: the 44 BFI-items). Second, the validities of those candidate scales are compared on a set of *a priori* defined validity guidelines (here: convergent validity—that is, near-perfect correlations with extant agency scales *and* a theory-consistent nomological net—that is, a negative correlation with social assimilation, etc.). Third, the AC method assigns a reward/weight to each item from the best performing candidate scale. Finally, Steps 1 to 3 are repeated for a predefined number of iterations (here: 40 iterations). Importantly, from the second iteration onward, the items for each candidate scale are no longer drawn at complete random from the pool of candidate items. Instead, items are more likely to be drawn if they were previously part of the best-performing candidate scales (i.e., if they received rewards/weights in previous iterations). When the number of iterations is sufficiently high, the last iteration's best performing scale promises to meet all validity guidelines (Armstrong et al., 1992; Yarkoni, 2010).

Strength and limitations. The AC method has two main strengths: (a) unlike the TS method, the AC method results in scales based on the performance of *candidate scales* rather than candidate items—a major advantage (see description of TS method). (b) The AC method allows scale selection based on more than one validity guideline. Still, the AC method also has two main limitations: (a) even though the AC method evaluates the performance of scales rather than the performance of single items, that method assigns rewards/weights to single items. It is possible, however, that an item only performed well, because it appeared together with a specific set of other items (Stanton, 2000). (b) The

AC method lacks clear rules regarding the ideal number of candidate scales per iteration and the ideal number of iterations. Both values could substantially affect the results (Olaru et al., 2015).⁴

Brute Force Method. The BF method (Russell et al., 2004; Stanton, 2000) involves three steps. First, it creates all possible scales from the candidate items (here: the 44 BFI items). Second, it compares the validity of those scales on a set of *a priori* defined validity guidelines (here: near-perfect correlations with extant agency scales *and* negative correlation with social assimilation *and* etc.). Finally, the BF method selects the scale that best meets those guidelines (Armstrong et al., 1992; Yarkoni, 2010).

Strengths and limitations. The BF method has two main strengths: (a) It necessarily yields the best possible scale (given the available items) because it takes into account all possible candidate scales. (b) Single items are not considered in isolation, but in concert with others—a major advantage (Stanton, 2000). The method also has two main limitations: (a) the BF method requires a great amount of computational power and time, because the number of candidate scales “explode” with an increasing number of candidate items (e.g., 20 items $\rightarrow \approx 180$ thousand possible 10-item scales; 40 items $\rightarrow \approx 850$ million possible 10-item scales). (b) Scale selection is based on a huge number of statistical analyses. Consequently, chance findings are practically certain. Fortunately, one can effectively address the latter limitation by testing for the robustness/replicability of BF results in multiple samples.

Present Research

The present research used three samples with a total sample size of 942. We relied on self-reports to construct the new BFI-Agency-Communion Scales for two reasons. First, most studies in personality psychology are self-report studies (Baumeister et al., 2007; Paulhus & Vazire, 2007). Second, most extant data sets containing the BFI rely on self-reports.

Each sample ($N_1 = 308$, $N_2 = 311$, $N_3 = 323$) exceeded the size required for correlations to stabilize (Schönbrodt & Perugini, 2013). We used the first two samples for scale construction. More precisely, we used the first sample to explore the suitability of the different scale constructions methods for extracting agency and communion scales from the BFI-items. Thus, we label the first sample “Exploratory Sample.” Because of its exploratory nature, it was crucial to follow up with a check for robustness of the results. We used the second sample for such a check and, accordingly, label it “Robustness Sample.” Based on the results of those two samples, we selected the final BFI-Agency-Communion Scales. Finally, to evaluate the replicability of newly

constructed scales, it is essential to draw on a sample that is independent from the samples used for scale construction. We used our third sample for such an independent evaluation and, accordingly, apply the label “Evaluation Sample.”

Method

Below, we report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The data for the present research is available at <https://madata.bib.uni-mannheim.de/id/eprint/363>. The measures section provides references to where all materials used in this article can be obtained from. The data analysis scripts can be obtained from the first author on request.

Participants

Amazon’s Mechanical Turk (MTurk) is the most widely used participant pool in basic psychological research (Anderson et al., 2018). It is therefore likely that many extant BFI-studies rely on MTurk. Hence, it made sense to use MTurk for our research. We recruited 942 participants who met the following three criteria: First, they had to be U.S. residents. Second, they had to possess worker approval ratings above 95%. Those two criteria ensure highest data quality, superseding other methods that safeguard data quality (e.g., attention-checks; Peer et al., 2014). Finally, they had to pass standard tests for careful responding (Meade & Craig, 2012; see online Supplement S1).

Participants completed the study in approximately 30 minutes and received US\$4.00 for full completion. After collecting the data, we randomly split them into three samples, thereby yielding the Exploratory Sample ($n = 308$, 49.5% female, $M_{age} = 33.92$ years, $SD_{age} = 10.88$), the Robustness Sample ($n = 323$, 52.6% female, $M_{age} = 33.22$ years, $SD_{age} = 11.62$), and the Evaluation Sample ($n = 311$, 52.9% female, $M_{age} = 34.13$ years, $SD_{age} = 10.64$).

Measures

This study was part of a larger project that included 11 measures. Here, we focus on the eight measures relevant to the present research (online Supplement S2 lists the three additional measures). We used the original, published response formats of all measures, except for the self-esteem measure. For that measure, we used the same 7-point rating scale used for the below described Big Two Inventory (BTI) and the social assimilation measure (1 = *does not describe me at all*, 7 = *describes me extremely well*). The measures were presented in random order.

Big Five Inventory. The 44 BFI items served as candidate items (John & Srivastava, 1999). The BFI measures each Big Five domain with 8 to 10 items. The five scales showed

high internal consistencies in all three samples (agreeableness: $.84 \leq \alpha_s \leq .87$; conscientiousness: $.89 \leq \alpha_s \leq .90$; extraversion: $.90 \leq \alpha_s \leq .91$; openness: $.86 \leq \alpha_s \leq .90$; neuroticism: $.91 \leq \alpha_s \leq .92$).

Extant Agency and Communion Scales. We sought to capture agency and communion in their fundamental, most basic, or broadest sense. To this end, we aggregated the agency (communion) scores of four extant agency (communion) scales. We selected those scales because they are (a) well-established and often-used, (b) particularly broad, and (c) complement each other concerning their item-content (extant scales focus on slightly different aspects of agency and communion). The result was a maximally broad *agency composite score* and a maximally broad *communion composite score*.⁵ The four extant scales were (a) Abele et al.’s (2008) Big Two trait adjectives (e.g., agency: “intelligent,” communion: “loyal”); (b) Saucier, Thalmeyer, Payne et al.’s (2014) Dynamism and Social Self-Regulation Scales (e.g., agency: “brave,” communion: “obedient”); (c) Spence and Helmreich’s (1978) Instrumental and Expressive Scales (e.g., agency: “Goes to pieces under pressure,” communion: “Has difficulty devoting self completely to others” (both reverse-keyed); and (d) Trapnell and Paulhus’s (2012) Agentic and Communal Values Scales (e.g., agency: “Wealth: financially successful, prosperous,” communion: “Equality: human rights and equal opportunity for all”). We modified instructions of the latter scale to assess traits, not values (“To what degree do you possess these attributes?” Maio, 2010). Across the three samples, all alphas were solid (agency: $.79 \leq \alpha_s \leq .89$, communion: $.84 \leq \alpha_s \leq .90$), as were the alphas of their composites (agency: $\alpha = .90$, communion: $\alpha = .92$).

In addition, we included the BTI (Gebauer, 2015). In contrast to the measures listed above, the BTI promises to assesses trait agency and trait communion in their full construct breadth (e.g., agency: “I am a competitive person, try to outperform others” and “I influence others’ lives;” communion: “I am very attentive to the needs of my loved ones” and “It takes a lot until I get angry at someone”). Alphas were strong across samples (agency: $.88 \leq \alpha_s \leq .89$, communion: all $\alpha_s = .91$).

Social Assimilation. Social assimilation (vs. contrast) is one of agency’s and communion’s defining covariates (see Introduction). Therefore, we included the 20-item Social Assimilation Scale (Gebauer, 2015). Sample items include “I follow social conventions and norms” and “I generally resist social pressure, don’t obey societal norms”). The scale showed strong alphas across the three samples ($.92 \leq \alpha_s \leq .93$).

Self-Esteem. Self-esteem is another defining covariate of agency and communion (see Introduction section).

Therefore, we included the 10-item Self-Esteem Scale (Rosenberg, 1965). Sample items include “I take a positive attitude toward myself” and “I feel I do not have much to be proud of” (reverse-keyed). Again, the alphas were strong across samples ($.94 \leq \alpha_s \leq .95$).

Results

This section has three parts. Part 1 describes the construction of our BFI-Agency-Communion Scales (i.e., results involving the Exploratory and Robustness Samples). Part 2 describes the empirical evaluation of those scales (i.e., results involving the Evaluation Sample). Part 3 compares the performance of the BFI-Agency-Communion Scales with that of extant agency and communion scales (including results involving the Evaluation Sample).

Part 1: Scale Construction

The Introduction section provided a brief overview of our four scale construction methods. We had to carefully adapt those methods to the specific task at hand. Here, we describe the task-specific adaptations of each method alongside their ensuing results.

Expert Rating Method. The ER method entailed three steps. First, three established authorities on agency and communion and on the Big Five judged all 44 BFI-items with regard to agentic and communal content. Those experts were Robert R. McCrae, Paul D. Trapnell, and Jerry S. Wiggins. The ratings were collected in 1994 as part of another (unfinished) project. They were requested via personal communication by one of the present authors. The raters were not given explicit definitions of agency and communion, but were asked to rely on their personal understanding of the constructs. To rate the items, the experts used 7-point rating scales (1 = *not at all suitable*; 7 = *very suitable*).

Second, we averaged the three agentic and communal ratings for each item, yielding an *agentic (communal) saturation score* per item. This averaging was justified due to high interexpert agreement (agentic saturation: $r = .86$, $z = 5.56$, $p < .001$, intraclass correlation: $r_{ic} = .94$; communal saturation: $r = .83$, $z = 5.29$, $p < .001$, intraclass correlation: $r_{ic} = .95$).

Finally, we averaged all BFI-items that received an agentic (communal) saturation score higher than 5 (on 7-point scales). This procedure resulted in the final *8-item ER-Agency Scale* and the final *9-item ER-Communion Scale*. Table 1 lists the items of the ER-Agency-Communion Scales. The scales showed solid alphas in the Exploratory and Robustness Samples, agency: $.82 \leq \alpha_s \leq .88$, communion: $.82 \leq \alpha_s \leq .83$. Alphas in the Evaluation Sample are presented in Part 2 of this Results section.

Target Scale Method. The TS method entailed four steps. First, we correlated each BFI-item with the agency composite and BTI-agency (hereafter: the two agency TSs). We inspected those correlations in both the Exploratory and Robustness Sample.

Second, we averaged the *five* BFI-items that evidenced the highest correlations with the two agency TSs across both samples (Exploratory and Robustness). The result was a five-item candidate TS-Agency Scale. We repeated that averaging-procedure until we obtained candidate TS-Agency Scales of lengths 5 to 14 items.

Third, we correlated all those candidate TS-Agency Scales with the two agency TSs. As it turned out, all candidate scales evidenced correlations of $r_{da} \geq .80$ (we use the subscript “da” to denote correlations disattenuated for measurement unreliability via Spearman-Brown correction).⁶ Most authorities agree that a disattenuated correlation of $r_{da} \geq .80$ signals indistinguishability of two constructs (Combs, 2010; Gray, 2017).

Hence, we considered an additional validity guideline to select our TS-Agency Scale. We chose the correlation of our candidate TS-Agency Scales with the two *communion* TSs (communion composite and BTI-communion). We expected our final TS-Agency Scale to correlate no more than $r_{da} \leq .40$ with both communion TSs because earlier literature suggested a disattenuated correlation of $r_{da} = .40$ as the upper limit for correlations between agency and communion (Abele et al., 2008: $r_{da} = .34$; Abele et al., 2016: $r_{da} = .40$; Wojciszke et al., 2011: $r_{da} = .30$). Only one candidate scale met this criterion, resulting in the final *eight-item TS-Agency Scale*.

To construct the TS-Communion Scale, we used analog steps as described in the previous paragraph. First, we correlated each BFI-item with the communion TSs. Second, we computed the candidate TS-Communion Scales (5-14 items). Third, we examined the disattenuated correlations of those candidate scales with the communion TSs and found all candidate scales to meet the validity guideline of $r_{da} \geq .80$.

Finally, we examined the disattenuated correlation of all candidate scales with the agency TSs: All candidate scales met the validity guideline of $r_{da} \leq .40$. Hence, we chose the 8-item candidate scale as our final *TS-Communion Scale*—simply to equate the lengths of the two TS-Agency-Communion Scales. Table 1 lists the items of the TS-Agency-Communion Scales. The two scales showed similarly strong reliabilities in the Exploratory and Robustness Samples (agency: both $\alpha_s = .88$, communion: $.83 \leq \alpha_s \leq .84$).

Ant Colony Method. The AC method uses a nature-informed optimization algorithm to compare the validity of different candidate scales (see Introduction section). Following recommendations by Leite et al. (2008), we compared 60

Table 1. BFI-Items Emerging from our Four Scale Construction Methods.

BFI Domains/Facet	Items	Agency	Communion
	I see myself as someone who . . .		
A-Compliance	. . . tends to find fault with others. ^a		TS, AC
A-Altruism	. . . is helpful and unselfish with others.		ER, TS, BF
A-Compliance	. . . starts quarrels with others. ^a		TS, AC, BF
A-Compliance	. . . has a forgiving nature.		ER, TS, BF
A-Altruism	. . . is generally trusting.		ER, AC
A-Altruism	. . . can be cold and aloof. ^a		BF
A-Altruism	. . . is considerate and kind to almost everyone.		ER, TS, AC, BF
A	. . . is sometimes rude to others. ^a		TS, AC, BF
A	. . . likes to cooperate with others.		ER, TS, AC, BF
C	. . . does a thorough job.		ER, AC, BF
C-Order	. . . can be somewhat careless. ^a		
C-Self Discipline	. . . is a reliable worker.		ER
C-Order	. . . tends to be disorganized. ^a		
C-Self Discipline	. . . tends to be lazy. ^a		
C-Self Discipline	. . . perseveres until the task is finished.	BF	ER
C	. . . does things efficiently.		TS
C-Self Discipline	. . . makes plans and follows through with them.		ER
C-Self-Discipline	. . . is easily distracted. ^a		
E-Assertiveness	. . . is talkative.	ER, TS	
E-Assertiveness	. . . is reserved. ^a	BF	
E-Activity	. . . is full of energy.	ER, TS, AC	
E-Activity	. . . generates a lot of enthusiasm.	ER, TS	
E-Assertiveness	. . . tends to be quiet. ^a	TS, BF	
E-Assertiveness	. . . has an assertive personality.	ER, TS, AC	
E-Assertiveness	. . . is sometimes shy, inhibited. ^a	TS, AC	
E	. . . is outgoing, sociable.	ER, TS	
O	. . . is original, comes up with new ideas.	ER, TS, BF	
O-Ideas	. . . is curious about many different things.	ER	
O-Ideas	. . . is ingenious, a deep thinker.	BF	
O	. . . has an active imagination.	AC	
O-Ideas	. . . is inventive.	ER, BF	
O-Aesthetics	. . . values artistic, esthetic experiences.		
O-Ideas	. . . prefers work that is routine. ^a	AC	
O-Ideas	. . . likes to reflect, play with ideas.		
O-Aesthetics	. . . has few artistic interests. ^a		
O-Aesthetics	. . . is sophisticated in art, music, or literature.	BF	
N-Depression	. . . is depressed, blue.		
N-Anxiety	. . . is relaxed, handles stress well.	BF	
N	. . . can be tense.		
N-Anxiety	. . . worries a lot.		
N	. . . is emotionally stable, not easily upset. ^a		
N-Depression	. . . can be moody.		
N-Anxiety	. . . remains calm in tense situations.	AC	
N-Anxiety	. . . gets nervous easily.		

Note. ER = expert rating method; TS = target scale method; AC = ant colony method; BF = brute force method; A = agreeableness; C = conscientiousness; E = extraversion; O = openness; N = neuroticism.
^aIndicates reversed keyed items.

different candidate scales in each of 40 iterations.⁷ We also adopted other validity guidelines (see Table 2, left-hand side). A narrative description of those guidelines follows:

Regarding *convergent validity*, we again aspired to a near-perfect correlation of our AC-Agency-Communion Scales with extant agency and communion scales. Similarly, our

Table 2. Validity Guidelines for the AC-Agency-Communion Scales (Left-Hand Side) and Part-2 Results in the Evaluation Sample (Right-Hand Side).

	Validity guidelines		ER		TS		AC		BF	
	agy	com	agy	com	agy	com	agy	com	agy	com
<i>Convergent and discriminant validity</i>										
	agy-comp: $r_{da} \geq .80$	com-comp: $r_{da} \geq .80$.83	.89	.83	.94	.99	.93	.82	.97
	BTI-agy: $r_{da} \geq .80$	BTI-com: $r_{da} \geq .80$.90	.89	.87	.97	1.00	.99	.94	.99
	com-comp: $.00 \leq r_{da} \leq .40$	agy-comp: $.00 \leq r_{da} \leq .40$.37	.56	.28	.45	.32	.43	.40	.44
	BTI-com: $.00 \leq r_{da} \leq .40$	BTI-agy: $.00 \leq r_{da} \leq .40$.42	.47	.35	.37	.47	.35	.50	.37
<i>External criteria</i>										
Social assimilation	$r_{da} \leq -.10$	$r_{da} \geq .10$	-.31	.16	-.23	.24	-.38	.28	-.37	.25
	$r_{com\ controlled} \leq -.30$	$r_{agy\ controlled} \geq .30$	-.42	.33	-.29	.29	-.41	.36	-.49	.46
Self-esteem	$.00 \leq r_{da} < .80$	$.00 \leq r_{da} < .80$.50	.42	.51	.53	.63	.51	.52	.52
	$r_{(agy, se)_{com\ controlled}} - r_{(com, se)_{agy\ controlled}} > .00$		-.04		.00		.08		.00	
<i>Internal structure (items as indicators)</i>										
	CFI $\geq .95$	CFI $\geq .95$.74	.71	.88	.96	.95	.96	.98	.97
	RMSEA $\leq .08$	RMSEA $\leq .08$.22	.19	.17	.08	.08	.07	.06	.06
	SRMR $\leq .08$	SRMR $\leq .08$.11	.12	.07	.04	.05	.04	.06	.04
<i>Internal structure (parcels as indicators)</i>										
	CFI $\geq .95$	CFI $\geq .95$.99		.98		.92		.91	
	RMSEA $\leq .08$	RMSEA $\leq .08$.07		.07		.13		.16	
	SRMR $\leq .08$	SRMR $\leq .08$.04		.04		.06		.08	
Interrelation	$r_{da}(agy, com) \leq .40$.51		.47		.43		.49	
Interrelation (acquiescence controlled)			.39		.30		.41		.39	
Reliability (Cronbach's α)			.86		.85		.89		.82	

Note. agy = agency; com = communion; comp = composite score; BTI = Big Two Inventory (Gebauer, 2015); r_{da} = correlation is corrected for measurement unreliability; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

expectation for *discriminant validity* was a moderate correlation of $r_{da} \leq .40$ between the AC-Agency Scale (AC-Communion Scale) and our two target communion (agency) scales. We did so, because associations of that size are typical among agency and communion (Abele et al., 2008: $r_{da} = .34$; Abele et al., 2016: $r_{da} = .40$; Wojciszke et al., 2011: $r_{da} = .30$).

Regarding *external criteria*, we chose two central criteria—social assimilation (Bakan, 1966) and self-esteem (Abele et al., 2016). Based on earlier literature, we anticipated a negative correlation between the AC-Agency Scale and social assimilation and a positive correlation between the AC-Communion Scale and social assimilation (Abele & Wojciszke, 2007; Gebauer et al., 2013). Furthermore, we aspired to positive, but nonperfect, correlations between the AC-Agency-Communion Scales and self-esteem: Of the two, we expected a stronger correlation with the AC-Agency Scale than with the AC-Communion Scale (Abele et al., 2016; Wojciszke et al., 2011).

Regarding the *internal structure* of each candidate AC-Agency-Communion Scale, we fit confirmatory factor analyses (CFAs) to the candidate scales. The items of each candidate scale served as indicators of a latent variable (agency or communion). Additionally, we specified a method factor that was indicated by all reverse-keyed

items (if there were any). We relied on conventional model fit criteria: CFA $\geq .95$, RMSEA $\leq .08$, and SRMR $\leq .08$. Finally, regarding the *interrelation* between the AC-Agency-Communion Scales, we aspired to an interrelation similar to extant agency and communion scales. Hence, we used $r_{da} \leq .40$.

We applied the AC method separately in the Exploratory and Robustness Samples and for scales of different lengths (5-14 items). Subsequently, we compared the best candidate AC-Agency-Communion Scales across the two samples. Because the single best AC-Agency-Communion Scale in the Exploratory Sample differed from the single best AC-Agency-Communion Scale in the Robustness Sample, we selected the AC-Agency-Communion Scale with the best properties on average across the two samples. This resulted in the final *six-item AC-Agency Scale* and the final *seven-item AC-Communion Scale*. Table 1 lists the items of the AC-Agency-Communion Scales. Considering the brevity of the AC-Agency Scale, both scales were sufficiently reliable in the Exploratory and Robustness Samples (agency: $.61 \leq \alpha \leq .64$, communion: $.78 \leq \alpha \leq .79$).

Brute Force Method. The BF method entailed three steps. First, we needed to narrow down the initial item pool, because it was computationally too taxing to consider all

candidate BF-Agency-Communion Scales (the 44 BFI-items result in a total of 199,045,391,241 candidate scales). Hence, we applied the following restrictions: The item pool for the BF-Agency (BF-Communion) Scale was restricted to items that correlated substantially, $r \geq .25$, with the agency composite (communion composite) and with BTI-agency (BTI-communion). These restrictions yielded 28 candidate items for the BF-Agency Scale (BF-Communion Scale).

Second, we generated an exhaustive candidate scale pool from the two sets of 28 items. We restricted the length of our candidate scales between 5 items and 14 items, resulting in a total of 308,503,740 candidate scales: All possible 5-item scales (98,280 candidate scales) to all possible 14-item scales (40,116,600 candidate scales). Finally, we compared all candidate scales to select the best BF-Agency-Communion Scales.

To select the best BF-Agency-Communion Scales, we applied the same validity guidelines as for the AC method (left-hand side of Table 2). Also as in the AC method, we applied those validity guidelines in both, the Exploratory and Robustness Samples. In total, 120 pairs of candidate BF-Agency and BF-Communion Scales met all validity guidelines across both samples. From those candidate scales, we chose agency and communion scales of equal length, yielding the final *eight-item BF-Agency Scale* and the final *eight-item BF-Communion Scale* (Table 1). Both scales were reliable in the Exploratory and Robustness Samples (agency: $.70 \leq \alpha \leq .73$, communion: $.80 \leq \alpha \leq .84$).

Part 2: Scale Evaluation

In this part, we evaluate the validity of our newly constructed scales based on the degree to which those scales meet the *validity guidelines* described in Part 1 (left-hand side of Table 2). To this end, we used the Evaluation Sample, which was entirely independent from the samples used to construct our scales (Exploratory and Robustness Samples).⁸

Expert Rating Method. Table 2 (right-hand side) describes the performance of all BFI-Agency-Communion Scales in meeting our *validity guidelines* (see Part 1 of the results section for a description of those guidelines). The first column describes the performance of the ER-Agency-Communion Scales. The upper-left panel of Figure 1 visualizes that performance graphically. The ER-Agency Scale met most of our validity guidelines (agy-comp: $r_{da} \geq .80$; BTI-agy: $r_{da} \geq .80$; com-comp: $.00 \leq r_{da} \leq .40$; social assimilation: $r_{da} \leq -.10$, $r_{com\ controlled} \leq -.30$; self-esteem: $.00 \leq r_{da} < .80$) and the scale was reliable ($\alpha = .86$). Likewise, the ER-Communion Scale also met most of our validity guidelines (com-comp: $r_{da} \geq .80$; BTI-com: $r_{da} \geq .80$; social assimilation: $r_{da} \geq .10$, $r_{agy\ controlled} \geq .30$; self-esteem: $.00 \leq r_{da} < .80$) and it was reliable, too ($\alpha = .85$).

Nonetheless, on a few occasions the ER-Agency-Communion Scales did not quite meet the validity guidelines. The ER-Agency Scale was somewhat too highly related to BTI-communion (validity guideline: $.00 \leq r_{da} \leq .40$, result: $r_{da} = .42$), its correlation with self-esteem was not higher than the ER-Communion Scale's correlation with self-esteem (validity guideline: $r_{(agy, se)_{com\ controlled}} - r_{(com, se)_{agy\ controlled}} > .00$, result: $r_{da} = -.04$), and the ER-Agency Scale's interrelation with the ER-Communion Scale was somewhat too high (validity guideline: $r_{da} \leq .40$, result: $r_{da} = .51$), but met expectations when we controlled for acquiescence ($r_{da} = .39$).⁹ Finally, a CFA evidenced poor fit, when items served as indicators of ER-agency (comparative fit index [CFI] = .74, root mean square error of approximation [RMSEA] = .22, standardized root mean square residual [SRMR] = .11), but the fit was good when parcels served as indicators in a CFA with ER-agency and ER-communion as correlated factors (CFI = .99, RMSEA = .07, SRMR = .04).¹⁰ Additionally, the ER-Communion Scale was somewhat too highly related to the agency composite (validity guideline: $.00 \leq r_{da} \leq .40$, result: $r_{da} = .56$) and to BTI-agency (validity guideline: $.00 \leq r_{da} \leq .40$, result: $r_{da} = .47$). Finally, a CFA evidenced poor fit, when items served as indicators of ER-communion (CFI = .71, RMSEA = .19, SRMR = .12), but the fit was good when item-parcels served as indicators in a CFA with ER-agency and ER-communion as correlated factors (see above). Overall, the ER-Agency-Communion Scales performed well, with two exceptions. First, we needed to control for acquiescence in order to meet the validity guideline on the interrelation between the two ER-Agency-Communion Scales. Second, we needed to parcel the items of the ER-Agency-Communion Scales in order to receive good model fit. The Discussion section describes why those two exceptions do not seriously question the validity of the ER-Agency-Communion Scales.

Target Scale Method. Table 2 also details the performance of the TS-Agency-Communion Scales (second column of right-hand side). The upper-right panel of Figure 1 illustrates that performance graphically. The TS-Agency Scale met most of our validity guidelines (agy-comp: $r_{da} \geq .80$; BTI-agy: $r_{da} \geq .80$; com-comp: $.00 \leq r_{da} \leq .40$; BTI-com: $.00 \leq r_{da} \leq .40$; social assimilation: $r_{da} \leq -.10$, self-esteem: $.00 \leq r_{da} < .80$) and the scale was reliable ($\alpha = .89$). Likewise, the TS-Communion Scale also met most of our validity guidelines (com-comp: $r_{da} \geq .80$; BTI-com: $r_{da} \geq .80$; BTI-agy: $.00 \leq r_{da} \leq .40$; social assimilation: $r_{da} \geq .10$, $r_{agy\ controlled} \geq .30$; self-esteem: $.00 \leq r_{da} < .80$; internal structure—items as indicators: CFI $\geq .95$, RMSEA $\leq .08$, SRMR $\leq .08$) and it was reliable, too ($\alpha = .82$).

On a few occasions the TS-Agency-Communion Scales did not quite meet the validity guidelines. The correlation of the TS-Agency Scale with social assimilation was not quite as negative as expected, when TS-Communion was

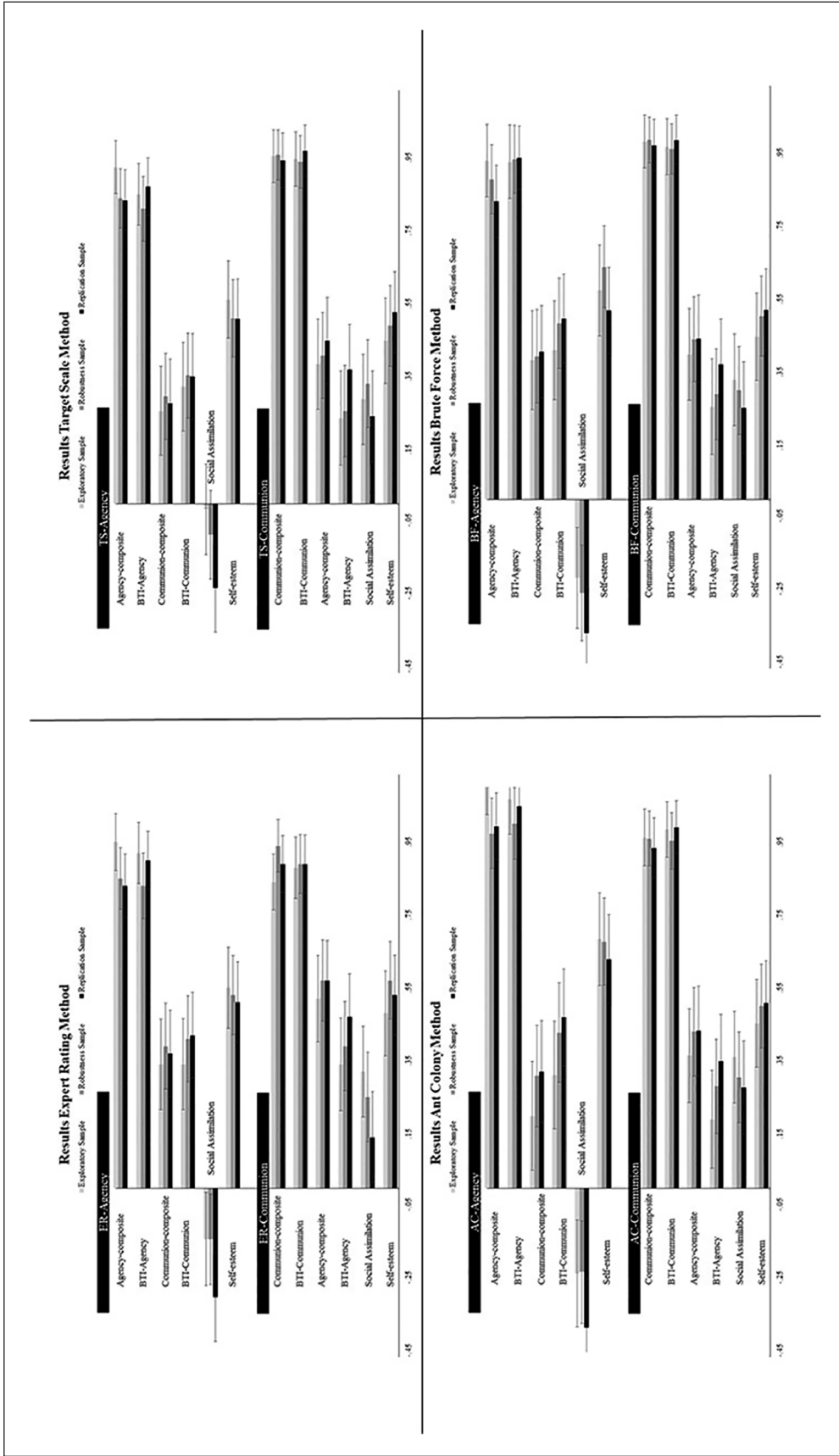


Figure 1. Disattenuated correlations of the BFI-Agency-Communism Scales. Note. BFI = Big Five Inventory; BTI = Big Two Inventory; ER = expert rating method; TS = target scale method; AC = ant colony method; BF = brute force method.

Table 3. Mean and Range of Correlations Between the Four BFI-Agency-Communion Scales and the Five Extant Agency and Communion Scales in the Evaluation Sample (Before Disattenuation).

	Mean correlations with . . .	
	. . . extant agency scales	. . . extant communion scales
ER-Agency	$\bar{r} = .68 (.54 \leq r \leq .79)$	ER-Communion $\bar{r} = .73 (.55 \leq r \leq .78)$
TS-Agency	$\bar{r} = .69 (.55 \leq r \leq .82)$	TS-Communion $\bar{r} = .77 (.59 \leq r \leq .84)$
AC-Agency	$\bar{r} = .72 (.64 \leq r \leq .81)$	AC-Communion $\bar{r} = .74 (.56 \leq r \leq .84)$
BF-Agency	$\bar{r} = .64 (.46 \leq r \leq .77)$	BF-Communion $\bar{r} = .77 (.61 \leq r \leq .85)$
Abele-Agency	$\bar{r} = .75 (.64 \leq r \leq .84)$	Abele-Communion $\bar{r} = .74 (.66 \leq r \leq .77)$
Saucier-Agency	$\bar{r} = .75 (.62 \leq r \leq .84)$	Saucier-Communion $\bar{r} = .74 (.63 \leq r \leq .79)$
Spence-Agency	$\bar{r} = .71 (.67 \leq r \leq .75)$	Spence-Communion $\bar{r} = .63 (.63 \leq r \leq .66)$
Trapnell-Agency	$\bar{r} = .66 (.62 \leq r \leq .69)$	Trapnell-Communion $\bar{r} = .74 (.63 \leq r \leq .78)$
BTI-Agency	$\bar{r} = .73 (.69 \leq r \leq .78)$	BTI-Communion $\bar{r} = .74 (.59 \leq r \leq .79)$

Note. For the four new agency and communion scales, the values reflect the mean associations between each new agency (communion) scale and all five extant agency (communion) scales. For the five extant agency and communion scales, the average correlations reflect the mean association between each extant agency (communion) scale and the four other extant agency (communion) scales. BFI = Big Five Inventory; ER = expert rating method; TS = target scale method; AC = ant colony method; BF = brute force method; Abele = Big Two trait adjectives (Abele et al., 2008); Saucier = Dynamism and Social Self-Regulation Scales (Saucier, Thalmeyer, Payne et al., 2014); Spence = Personal Attributes Questionnaire (Spence & Helmreich, 1978); Trapnell = Agentic and Communal Values Scale (Trapnell & Paulhus, 2012); BTI = Big Two Inventory (Gebauer, 2015).

controlled (validity guideline: $r_{com\ controlled} \leq -.30$, result: $r_{da} = -.29$), the correlation of the TS-Agency Scale with self-esteem was not higher than the correlation of the TS-Communion Scale with self-esteem (validity guideline: $r[agy, se]_{com\ controlled} - r[com, se]_{agy\ controlled} > .00$, result: $r_{da} = .00$), and the interrelation between the two TS-Agency-Communion Scales was somewhat too high (validity guideline: $r_{da} \leq .40$, result: $r_{da} = .47$), but met expectations when we controlled for acquiescence ($r_{da} = .30$). Finally, a CFA evidenced poor fit, when items served as indicators of TS-agency (CFI = .88, RMSEA = .17, SRMR = .07), but the fit was good when parcels served as indicators in a CFA with TS-agency and TS-communion as correlated factors (CFI = .98, RMSEA = .07, SRMR = .04). Furthermore, the TS-Communion Scale was somewhat too highly related to the agency composite (validity guideline: $.00 \leq r_{da} \leq .40$, result: $r_{da} = .45$).

Overall, the TS-Agency-Communion Scales performed well, with the same two exceptions as in the ER-Agency-Communion Scales (their intercorrelation when acquiescence was not controlled and model fit when items served as indicators). The Discussion section describes why those expectations do not seriously question the validity of the TS-Agency-Communion Scales.

Ant Colony Method. Table 2 also details the performance of the AC-Agency-Communion Scales (third column of right-hand side). The lower-left panel of Figure 1 illustrates that performance graphically. The AC-Agency Scale met most of our validity guidelines (agy-comp: $r_{da} \geq .80$; BTI-agy: $r_{da} \geq .80$; com-comp: $.00 \leq r_{da} \leq .40$; social assimilation: $r_{da} \leq -.10$, $r_{com\ controlled} \leq -.30$; self-esteem: $.00 \leq r_{da} < .80$; $r[agy, se]_{com\ controlled} - r[com, se]_{agy\ controlled} > .00$; internal

structure—items as indicators: CFI $\geq .95$, RMSEA $\leq .08$, SRMR $\leq .08$) and the scale was sufficiently reliable for a six-item scale of the broad agency trait ($\alpha = .67$). Similarly, the AC-Communion Scale met our validity guidelines (with one exception) and the scale was reliable ($\alpha = .79$). Again, on a few occasions the AC-Agency-Communion Scales did not quite meet the validity guidelines. The AC-Agency Scale was somewhat too highly related to BTI-communion (validity guideline: $.00 \leq r_{da} \leq .40$, result: $r_{da} = .47$) and its interrelation with the AC-Communion Scale was somewhat too high (validity guideline: $r_{da} \leq .40$, result: $r_{da} = .43$). As to the AC-Communion Scale, only its interrelation with the AC-Agency Scale was somewhat too high (see above).

Overall, the AC-Agency-Communion Scales performed well, with one divergence from our validity guidelines only—their interrelation was a bit higher than we (initially) expected.

Brute Force Method. Table 2 also describes the performance of the BF-Agency-Communion Scales (final column of right-hand side). The lower-right panel of Figure 1 illustrates that performance graphically. The BF-Agency Scale met most of our validity guidelines (agy-comp: $r_{da} \geq .80$; BTI-agy: $r_{da} \geq .80$; com-comp: $.00 \leq r_{da} \leq .40$; social assimilation: $r_{da} \leq -.10$, $r_{com\ controlled} \leq -.30$; self-esteem: $.00 \leq r_{da} < .80$; internal structure—items as indicators: CFI $\geq .95$, RMSEA $\leq .08$, SRMR $\leq .08$) and the scale was reliable ($\alpha = .76$). Likewise, the BF-Communion Scale also met most of our validity guidelines (com-comp: $r_{da} \geq .80$; BTI-com: $r_{da} \geq .80$; BTI-agy: $.00 \leq r_{da} \leq .40$; social assimilation: $r_{da} \geq .10$, $r_{agy\ controlled} \geq .30$; self-esteem: $.00 \leq r_{da} < .80$; internal structure—items as indicators: CFI $\geq .95$, RMSEA $\leq .08$, SRMR $\leq .08$) and it was also reliable, too ($\alpha = .82$).

On a couple of occasions, the BF-Agency-Communion Scales appeared to flout the validity guidelines. The BF-Agency Scale was somewhat too highly related to BTI-communion (validity guideline: $.00 \leq r_{da} \leq .40$, result: $r_{da} = .50$) and its inter-relation with the BF-Communion Scale was somewhat too high (validity guideline: $r_{da} \leq .40$, result: $r_{da} = .49$); but these values met expectations when we controlled for acquiescence ($r_{da} = .39$). Finally, the BF-Communion Scale was somewhat too highly related to the agency composite (validity guideline: $.00 \leq r_{da} \leq .40$, result: $r_{da} = .44$).

Overall, however, the BF-Agency-Communion Scales performed well, with one exception (their intercorrelation when acquiescence was not controlled). The Discussion section explains why this expectation does not seriously question the validity of the BF-Agency-Communion Scales.

Part 3: Comparison With Extant Agency-Communion Scales

The four BFI-Agency-Communion Scales show remarkable convergence. The mean interrelation between the four new agency scales was $\bar{r} = .86$ ($.79 \leq rs \leq .87$) and the mean interrelation between the four new communion scales was $\bar{r} = .90$ ($.85 \leq rs \leq .96$). Moreover, after disattenuating those correlations for unreliability, they all approached $r_{da} = 1.00$.¹¹

The four BFI-Agency-Communion Scales also showed remarkable convergence with extant agency and communion scales (Table 3). The mean interrelation between the four new agency scales was comparable to the mean interrelation between the five extant agency scales, $\bar{r} = .86$ vs. $\bar{r} = .72$, and the mean inter-relation between the four new communion scales was comparable to the mean inter-relation between the five extant communion scales, $\bar{r} = .90$ vs. $\bar{r} = .72$.

Most important, Table 3 also shows that each of the four new agency scales correlated about as highly with the five extant agency scales as those five extant scales correlated with each other, $\bar{r} = .68$ versus $\bar{r} = .72$. Likewise, each of the four new communion scales correlated about as highly with the five extant communion scales as those five communion scales correlated with each other, $\bar{r} = .75$ versus $\bar{r} = .72$. In other words, the performance of our BFI-Agency-Communion Scales was not out of line with that of extant scales; instead they appear to belong to the same family of scales (online Supplement S3 shows the full correlation matrix).

In another respect, too, the four BFI-Agency-Communion Scales fit comfortably within the norms of extant agency and communion scales. Specifically, online Supplement S4 lists the performance of the five extant scales in regard to our validity guidelines. This online supplement reveals that the extant scales generally meet our validity guidelines.

Notably, though, they do not meet those guidelines any more closely than the BFI-Agency-Communion Scales. Thus, the extant agency and communion scales and the BFI-Agency-Communion Scales appear similarly capable of capturing the two fundamental dimensions of agency and communion.

Discussion

At the *empirical* level, the present research set out to extract the two fundamental personality traits—agency and communion—from the BFI (John & Srivastava, 1999), the most popular (nonproprietary) measure of the Big Five. We approached that extraction with four alternative scale construction methods: ER, TS, AC, and BF. At the *conceptual* level, we sought to make two contributions: (a) we sought to extend knowledge about the relative efficacy of the four methods and (b) we sought to enable researchers to test agency and communion hypotheses with extant BFI data sets (data sets collected in their own labs and openly accessible, large-scale ones). As detailed next, we believe that we succeeded in both endeavors.

Relative Efficacy of the Four Alternative Scale-Construction Methods

The present research is the first to compare the four scale-construction methods head-to-head and it does so in the spirit of Hase and Goldberg (1967). In short, the four methods yielded highly similar results. Considering the diversity of those methods, this convergence is both reassuring and impressive. Nonetheless, application of the four scale-construction methods did yield some differences. We briefly discuss them next.

Expert Rating Method. Our application of the ER method used the perspectives of three established authorities (Robert R. McCrae, Paul D. Trapnell, and Jerry S. Wiggins). Their personal conceptualization of agency and communion were sufficiently similar to justify a composite ER. This high level of convergence indicates substantial agreement across (at least three) experts about the meaning of agency and communion. The resulting ER-Agency-Communion Scales had adequate psychometric properties and evidenced excellent convergent and discriminant validity. These results dispel concerns—at least regarding agency and communion—that the ER method yields scales with low reliability or low construct breadth (Stanton, 2000). In addition, the ER-Agency-Communion Scales correlated as expected with social assimilation and self-esteem—support for the scales' construct validity.

We also found, however, that the interrelation between the ER-Agency-Communion Scales was higher than originally expected, $r_{da} = .51$. Yet we noticed post hoc that an

interrelation of that size is not unusual (Buchanan & Bardi, 2015: $r_{da} = .58$; Gebauer et al., 2014: $r_{da} = .42$; Suitner & Maass, 2008: $r_{da} = .56$). In addition, after controlling for acquiescence, that interrelation dropped to $r_{da} = .39$. A likely reason is that the ER-Agency-Communion Scales do not contain reverse-keyed items. As a result, acquiescence inflated the interrelation of the two scales. Thus, we recommend controlling for acquiescence in future use of the ER-Agency-Communion Scales.

We also found poor model fit in a CFA that used the items of the ER-Agency Scale (ER-Communion Scale) as indicators of agency (communion). It is peculiar that the ER-Agency-Communion Scales evidenced such poor fit even though they passed other validity guidelines with flying colors. Yet that peculiarity may be easily reconciled. Agency and communion are considered the broadest content dimensions in psychology, much broader than the specific BFI-items (Abele & Wojciszke, 2007; Paulhus & John, 1998). Under such circumstances (construct-breadth > item-breadth), the reflective construct-to-item relation that theoretically underpins CFA models may be inappropriate (McCrae, 2015). One solution is to form item-parcels that are similarly broad than the construct assessed by those parcels (Little et al., 2002). We ran a CFA on the parcel-based model. The resultant model fit was satisfying (Table 2).

Target Scale Method. The TS method has few precedents. Undoubtedly most recognizable is the IPIP, where the TS method was used to develop nonproprietary substitutes for the proprietary NEO-PI-R scales (Goldberg et al., 2006; Johnson, 2014). Our TS-Agency-Communion Scales evidenced excellent convergence with extant scales. Somewhat unexpected was that the TS-Agency-Communion Scales correlated only modestly with our external criteria: social assimilation and self-esteem. Note that the TS method uses a single criterion for item-selection. Thus, additional theoretical assumptions for the final scales (e.g., associations with external criteria) are not automatically satisfied. Also unexpected were the high interrelation between the TS-Agency-Communion Scales and the poor model fit of the TS-Agency Scale. However, controlling for acquiescence reduced the interrelation and item-parceling improved model fit. In short, those unexpected results do not seriously question the validity of the TS-Agency-Communion Scales.

Ant Colony Method. Although it has served other purposes (Dorigo & Gambardella, 1997), the AC method has rarely been used in personality item selection and has undergone only one competition with other scale construction methods (Olaru et al., 2015). Our results confirmed that the AC method is effective as it yielded valid agency and communion scales. Those scales had satisfying convergent and discriminant validities, correlated as intended with our external criteria, showed a satisfactory model fit, and possessed low

interrelations with each other. Yet our results also revealed a limitation of the AC method. Specifically, the results were somewhat more variable than the results of other methods (i.e., the best scales differed somewhat across the Exploratory and Robustness Samples). We addressed that limitation by selecting scales that performed best across the two samples. This two-step selection strategy proved efficient and we can recommend it for future research.

Brute Force Method. The BF method also yielded valid agency-communion scales. The BF-Agency-Communion Scales had satisfying convergent and discriminant validities, correlated as intended with our external criteria, showed a satisfactory model fit, and were appropriately interrelated (after controlling for acquiescence). In at least one way, the method is superior to the others: It alone provides an exhaustive search of all item combinations (i.e., candidate scales). Yet it also has a downside: the method requires massive computational time and power. In our case, the computational power needed for the method exceeded the computing capacity of available computers. Thus, we had to reduce the initial item pool by dropping 16 items with a very low a priori likelihood to become part of the final scale (item-correlations of $r < .25$ with extant agency and communion scales). Nonetheless, the psychometric properties of the BF-Agency-Communion Scales are as desirable as the properties of our other three BFI-Agency-Communion Scales.

Broader Insights Into Scale Construction

We gained several broader insights into scale construction. For example, our scale construction efforts employed three samples of roughly 300 participants each. Our choice of that number and size of samples proved sufficient for the construction of agency and communion scales with very consistent/stable properties across samples: very similar correlations with extant agency and communion scales, very similar correlations with external criteria, and very similar interrelations between agency and communion (Figure 1). Scale-properties across samples were even highly consistent/stable for scales constructed with the AC and BF methods. This is noteworthy, because those two methods draw on extremely large numbers of statistical analyses, which invite nonreplicable chance findings. Evidently, the number and size of our samples were sufficiently large to counteract nonreplicable chance findings (see also Schönbrodt & Perugini, 2013). Thus, for future scale-construction efforts we can recommend use of three separate samples with about 300 participants each.

Furthermore, Amazon's MTurk is the single largest research pool for personality and social psychologists (Anderson et al., 2018). Therefore, we suspect that there are hundreds of MTurk-based BFI studies awaiting re-analysis with a focus on agency and communion. Although the

quality of data from MTurk is sometimes criticized, our data proved to be of high quality. More precisely, across three samples, we identified fewer careless responders than in typical student samples (see online Supplement S1).

On the Relation Between the Two Fundamental Dimensions and the Big Five

We gained three broader insights into the relation between the two fundamental dimensions and the Big Five. First, our results show that it is possible to create agency and communion scales based on items formerly developed to assess the Big Five. To the extent that agency and communion are exhaustive content dimensions at their (most basic) level of abstraction (Abele & Wojciszke, 2018; Wiggins, 2003), our results fortify the assumption that the Big Five are exhaustive, too (Goldberg, 1990; McCrae & John, 1992).

Second, we also found that the BFI-items more easily allowed the construction of communion scales than agency scales. One explanation for this finding is the relative saturation of agency and communion in the BFI-items. Although each Big Five domain contains both agentic and communal content (Wiggins & Trapnell, 1996), that content is typically arbitrarily weighted in Big Five measures. Apparently, BFI items place less emphasis on agency than on communion.¹²

Finally, there is strong agreement that (elements of) extraversion and agreeableness are particularly central to agency and communion, respectively (Gebauer et al., 2015; Paulhus & John, 1998; Wiggins, 1991).¹³ Our validity guidelines did not incorporate those features of agency and communion. Nonetheless, most items of our BFI-Agency Scales turned out to be extraversion items (56% on average, ranging from 25% in the BF-Agency Scale to 88% in the TS-Agency Scale). Likewise, most items of our BFI-Communion Scales turned out to be agreeableness items (79% on average, ranging from 56% in the ER-Communion Scale to 88% in the BF-/TS-Communion Scales). Those results buttress the central role of extraversion and agreeableness in agency and communion, respectively. Those results also demonstrate, however, that our BFI-Agency-Communion Scales are more than extraversion and agreeableness, much in line with our broad conceptualization of agency and communion (see Footnote 1).

On the Relation Between the Two Fundamental Dimensions and the Higher Order Factors of the Big Five

Factor analyses of the Big Five domains (not their items) revealed two higher order factors: alpha and beta (Digman, 1997). Alpha manifests in agreeableness, conscientiousness, and neuroticism; beta manifests in extraversion and openness (see also DeYoung, 2006). Is alpha (beta) a conceptual equivalent to communion (agency)? The literature provides three competing answers to this question. Probably

the most common answer is that the higher order factors are indeed conceptual equivalents of agency and communion (Digman, 1997; Hopwood et al., 2011; Srivastava et al., 2010). Another common answer is that the higher order factors are distant relatives of agency and communion—they are moderately related, but distinct (DeYoung et al., 2013; Rau et al., 2019). Finally, there is a middle-ground position (Campbell et al., 2002; Gebauer et al., 2012; Paulhus & John, 1998). That position states that agency and beta are close relatives and, thus, strongly related, albeit not identical. Communion and alpha, by contrast, are distant relatives and, thus, moderately related. The middle-ground position further states that a closer relative of communion is a higher order factor that manifests in agreeableness and conscientiousness, but not in neuroticism (the reason being that communion is a content dimension and neuroticism is a purely evaluative, content-free domain; Furr & Funder, 1998; Gebauer et al., 2015).

Our data allowed us to compare the three competing answers regarding the conceptual relation between the higher order factors and the two fundamental dimensions of agency and communion. To this end, we tested whether the higher order factors meet our validity guidelines to the same extent than our BFI-Agency-Communion Scales (see Table 2 for those guidelines). Online Supplement S4 presents the results. They revealed that the higher order factors failed eight of our validity guidelines. Moreover, in some critical cases the degree of failure was severe (our BFI-Agency-Communion Scales also failed some validity guidelines, but the degree of failure was never severe). For example, alpha (supposedly communion) was very strongly related to the agency composite, $r_{da} = .74$. Alpha was also much more strongly related to self-esteem than was beta, $\Delta r = -.53$. Thus, our data clearly contradict the view that the higher order factors and the fundamental dimensions of agency and communion are conceptual equivalents. Yet our data revealed strong support for the middle-ground position. More precisely, the data suggested that beta and agency are close conceptual relatives, but the two were nonidentical. For example, the correlation between beta and the agency composite was $r_{da} = .71$ (see online Supplement S4). At the same time, the higher order factor of agreeableness and conscientiousness (neuroticism not included) was a close conceptual relative to communion, but the two were also nonidentical (see online Supplement S4). Those ancillary results help elucidate a much debated question in personality psychology. Also, they clearly illustrate that the higher order factors are no suitable alternative to our BFI-Agency-Communion Scales.

On the Interrelation Between Agency and Communion

In line with extant research (Buchanan & Bardi, 2015; Gebauer et al., 2014; Suitner & Maass, 2008), agency was

positively intercorrelated with communion in our data. This was the case for all extant agency and communion scales and also for all BFI-Agency-Communion Scales (Table 2). Controlling for acquiescence reduced those intercorrelations, but they remained sizable. That result is suggestive of a general factor of personality (Musek, 2007; Van der Linden et al., 2010). One popular view describes this general factor as a halo-like, positive self-evaluation devoid of any semantic content (Anusic et al., 2009; Pettersson et al., 2012). Global self-esteem has been described as a proxy for that halo-like, positive self-evaluation (Anusic et al., 2009; Şimşek, 2012). Consequently, controlling for self-esteem should reduce the correlation between agency and communion (it should not nullify the correlation altogether, because self-esteem is a proxy, not a perfect measure of the halo-like, positive self-evaluation; Gebauer et al., 2013; Wojciszke et al., 2011). And, in fact, controlling for self-esteem substantially reduced the correlations between the BFI-Agency and BFI-Communion Scales (ER: .39 → .20, TS: .30 → .13, AC: .41 → .09, BF: .39 → .10).

Facilitation of Agency and Communion Research With Extant BFI-Data Sets

Inside and outside of psychology laboratories, the BFI has been among the most popular measures for assessing personality. Accordingly, there exist a great number of data sets that contain BFI data. By revisiting those data sets, researchers can now compile agency and communion scores on the basis of BFI-items and test novel agency and communion predictions. Researchers can also evaluate agency and communion predictions in publicly available large-scale archival data sets. Those large, high-quality data sets have become a major asset in psychology. Yet those data sets rarely contain measures of agency and communion. If those data sets include the BFI, however, our new measures can unlock further research on agency and communion.

Where does that leave researchers who used the revised version, the BFI-2 (Soto & John, 2017)? We have focused on the original version because it has been used in the large majority of existing data sets, some of which go back 30 years. Nonetheless, the items of the BFI and the BFI-2 overlap to a large degree. To address the overlap issue post hoc, we conducted an item comparison and selected the BFI-2-items that best represent the BFI-items. In most cases, the items of the BFI were identical or only slightly revised in the BFI-2. In those rare cases, however, in which there were no equivalent items available, we substituted those items that most resembled the original BFI-items. Online Supplement S5 provides our recommendations for scoring agency and communion scales from BFI-2 data.

Recommendation on the Use of the BFI-Agency-Communion Scales in Future Research

Our four different scale-construction methods yielded four different BFI-Agency-Communion Scales (with relatively little item-overlap; see Footnote 11). All those scales appear valid and no scale appears clearly superior to its alternatives. This renders it difficult to recommend any specific scale. Instead, we would rather like to recommend that future research tests their substantive research questions with all scales. Two benefits would ensue.

First, conceptually identical results with all scales would boost confidence in those results. Stated differently, use of all scales may serve as a robustness check. Robustness checks enjoy increasing popularity in psychology. Therefore, it is a strength that the present research yielded several valid scales (not a single one). This is especially the case since online supplements have become commonplace, allowing researchers to report the results of one BFI-Agency-Communion Scale in the main text and the results of the other scales in an online supplement.

Second, Cronbach and Meehl's (1955) classic "bootstraps effect" stipulates that the development of substantive theories and the development of scales to test those theories should progress in iterative circles. From that perspective, future research may unearth some differences in the properties of the four BFI-Agency-Communion Scales and those differences may stimulate theory development, which—in turn—may favor one scale over its alternatives.

Additionally, future research may provide new opportunities to compare the validity of the four BFI-Agency-Communion Scales. In that spirit, one recent article compared our BFI-Agency-Communion Scales regarding their degree of measurement invariance across 102 countries ($N = 2,672,820$; Gebauer et al., 2020). Strongest evidence for measurement invariance emerged for the AC- and BF-Agency-Communion Scales. This is perhaps unsurprising, because our AC- and BF-methods were the only ones that attended to structural properties during scale construction. Nonetheless, if measurement invariance across cultures is important for a future project, the AC- and BF-Agency-Communion Scales are particularly appropriate (for cautionary notes on the (over-)interpretation of poor measurement invariance indices, see Funder, 2020; Gebauer et al., 2020, Section S4 of their online supplement).

Limitations and Future Directions

Although the BFI-Agency-Communion Scales evidenced solid validity in our research, we recommend several avenues for future research to further buttress their validity. Off the top, future research should extend the criteria beyond self-report. Replication of the present patterns using informant-reports and behavioral outcomes would further

increase confidence in the new scales. Similarly, confidence in the new scales would also increase, if those scales revealed results consistent with agency and communion theory. In fact, one recent set of studies has already used the new scales and found results consistent with agency and communion theory (Gebauer et al., 2020).

The present research relied on U.S. MTurkers (for good reason, see Introduction section) and Gebauer et al. (2020) relied on the Gosling–Potter Internet Personality Project (Gosling et al., 2004). Thus, all research to date on the BFI-Agency-Communion Scales has been web-based. As a complement, future research should also rely on other populations. Perhaps the most interesting ones are indigenous populations, because they typically evidence no Big Five structure of personality, but an agency-communion structure (Saucier, Thalmeyer, Payne et al., 2014; Thalmayer et al., 2020). It would, thus, be interesting to examine the properties of the BFI-Agency-Communion Scales within indigenous societies and to explore whether those scales may actually fare better than the five original BFI domains. Gurven et al. (2013) conducted research relevant to this issue. Those scholars administered the BFI in a sample of forager-farmers in the Bolivian Amazon. The BFI did not reveal the Big Five structure typically found in Western, educated, industrialized, rich, and democratic societies (Henrich et al., 2010). Instead, exploratory factor analyses revealed a two-factor structure, which reasonably resembled agency and communion (see also Gurven et al., 2014; von Rueden et al., 2015).

Another avenue for future research concerns gender differences in agency and communion. The traditional finding is that agency is higher among men than women, whereas communion is higher among women than men (Bem, 1974). However, that traditional pattern has shifted over time. About 20 years ago, agency began to emerge equally high among men and women in Western countries. Communion, by contrast, has remained higher among women than among men (Twenge, 1997). It is unclear where we now stand. Given the trend, agency may have become higher among women than men—at least in some cultures. Conversely, communion may have become equally high among men and women. Our data suggests that the trend has stabilized (see online Supplement S6). Across our three samples, agency was equally high among men and women. Communion, by contrast, was higher among women than among men. In other words, the present research replicated Twenge's (1997) results from over 20 years ago. Notably, we obtained those results with the extant agency and communion scales and also with our BFI-Agency-Communion Scales. The results' similarity across the two sets of scales further buttresses the validity of the BFI-Agency-Communion Scales. The results from our samples notwithstanding, results may be different in large-scale archival data sets that allow the analysis of diverse subgroups, such

as different age groups and different countries with varying gender-equality scores. Our BFI-Agency-Communion Scales offer novel and promising opportunities to examine gender differences regarding agency and communion in large-scale, cross-cultural data sets that include the BFI.

Finally, in the present research, we followed a research tradition that conceptualizes agency and communion as particularly basic and broad dimensions—even more basic and broader than the Big Five (Abele & Wojciszke, 2018; Bakan, 1966; Wiggins, 2003). To be consistent with that conceptualization, we measured agency and communion with the following five extant scales: the Big Two trait adjectives (Abele et al., 2008), the Dynamism and Social Self-Regulation Scales (Saucier, 2014), the Personal Attributes Questionnaire (Spence & Helmreich, 1978), the Agentic and Communal Values Scale (Trapnell & Paulhus, 2012), and the BTI (Gebauer, 2015). The BFI-Agency-Communion Scales evidenced high convergent validity in relation to all five of those extant scales (see Table 3). Future research, however, should evaluate associations with other extant measures of agency and communion. First and foremost, future research should evaluate associations with measures from the interpersonal circumplex tradition. Those measures focus on interpersonal aspects of agency and communion (Markey & Markey, 2009), including interpersonal problems (Alden et al., 1990) and interpersonal values (Locke, 2000). Thus, those measures are narrower in their conceptualization of agency and communion than are the BFI-Agency-Communion Scales (cf. Abele & Wojciszke, 2018; Bakan, 1966; McCrae & Costa, 1989; Wiggins, 1991). Irrespective, it would be interesting to know how strongly the BFI-Agency-Communion Scales overlap with measures such as the Circumplex Scales for the Inventory of Interpersonal Problems (Alden et al., 1990), the IPIP-Interpersonal Circumplex Scales (Markey & Markey, 2009), the Revised Interpersonal Adjective Scales (Wiggins et al., 1988), and the Circumplex Scales of Interpersonal Values (Locke, 2000).

Conclusion

The present research used four alternative scale-construction methods to extract the two fundamental dimensions of agency and communion from the BFI—the most popular (nonproprietary) Big Five measure. Two main contributions ensued. First, the present research deepened our knowledge regarding the four alternative scale-construction methods in the realm of personality. It was reassuring—impressive even—that all four methods yielded suitable and highly similar agency and communion scales. Therefore, we can recommend all four methods for the construction of personality measures even under adverse conditions such as ours (i.e., a very limited item pool of 44 candidate items). Second, the present research allows

scholars to examine agency and communion hypotheses with extant BFI datasets, including (a) those previously collected in their own labs and (b) those accessible from large-scale archives. Of course, we personally hope that our BFI-Agency-Communion Scales will be taken up by others and we look forward to novel insights that those scales promise to unearth.

Declaration of Conflicting Interests

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

Funding

The author(s) disclosed the following financial support for the research, authorship and/or publication of this article: We acknowledge support from the German Research Foundation (DFG; Grants GE 2515/3-1 and GE 2515/6-1) and the Social Science and Humanities Council of Canada (435-2015-0417).

ORCID iD

Theresa M. Entringer  <https://orcid.org/0000-0003-1149-7852>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. The present work portrays the fundamental dimensions of agency and communion as the *most* basic or *broadest* content dimensions in psychology. As such, the present work follows a large and growing (or even booming) tradition in many realms of psychology (Abele & Wojciszke, 2018; Bakan, 1966; Fiske et al., 2007; Frimer et al., 2011; Gebauer et al., 2013; Hopwood et al., 2011; Imhoff & Koch, 2017; Martin & Slepian, 2020; McAdams et al., 1996; Paulhus & John, 1998; Saucier, Thalmayer, Payne, et al., 2014; Srivastava et al., 2010; Trapnell & Paulhus, 2012; Wiggins, 2003). Notably, though, there also exists a somewhat different conceptualization of agency and communion (Leary, 1957; McCrae & Costa, 1989). According to that conceptualization, agency and communion are purely interpersonal in nature and, thus, somewhat narrower (at about the same level of breadth than the Big Five). Among other things, we use the term “fundamental dimensions of agency and communion” to signal that this article follows the former, particularly broad conceptualization of agency and communion.
2. We coined this term after consulting with several scale-construction experts (Lee Anna Clark, John Johnson, Lewis Goldberg, and David Watson).
3. The IPIP homepage gives concrete instructions for IPIP-scale construction: “1) Correlate all of the IPIP items available at that time (e.g., . . .) with the external criterion (e.g., . . .), and rank order the IPIP items by the absolute value of those correlations. 1a). When developing IPIP scales targeted on constructs from a multiscale inventory, correlate all IPIP items available at that time with each of the inventory scales, categorize each IPIP item by the scale with which it has its highest correlation, and rank order the IPIP items within each of the resulting categories by the size of those correlations. This will ensure that all of the IPIP items selected for an IPIP scale correlate more highly with its criterion scale than with any of the others” (<https://ipip.ori.org/newScaleConstruction.htm>).
4. The AC method compares a large number of candidate scales (here: 2,400) regarding their performance on multiple validity guidelines (here: 10). Thus, scale selection is based on a large number of statistical analyses (here: 24,000 per trait). Hence, chance findings are almost certain. Fortunately, one can effectively address this limitation by testing for the robustness/replicability of AC results in multiple samples.
5. Prior to aggregating the four extant agency and communion scales, we tested for their shared variance. Following Krizan and Herlache (2018; see also Gebauer et al., 2013), we ran an EFA (oblimin rotation) that included the four agency and communion scales’ scores. The EFA revealed a very clear two-factor structure (eigenvalues: Factor 1 = 4.11, Factor 2 = 2.25, Factor 3 = 0.49). The four communion scales loaded strongly on the first factor (factor loadings > .82) and the four agency scales loaded strongly on the second factor (factor loadings > .79). There were no noteworthy cross-loadings between the two factors (factor loadings < .19).
6. We relied on α (Cronbach, 1951) to estimate measurement unreliability, because that estimate remains the most popular one. Notably, though, Hunt and Bentler (2015) showed that a variant of Guttman’s (1945) λ_4 —that is, $\lambda_{4(0.50)}$ —is a most accurate estimate of measurement unreliability, if the items of a scale are narrower than the construct measured by that scale (which is almost certainly the case in our research). Therefore, we recalculated all disattenuated correlations from our Results section’s Part 2, using $\lambda_{4(0.50)}$ instead of α . Those disattenuated correlations were smaller, but the difference was negligible: On average, those disattenuated correlations differed from the disattenuated correlations reported in the main text by $\Delta r_{da} = .01$ and, with one exception only, the validity guidelines were met to the same (high) degree than in the main text (i.e., when α was used to estimate measurement unreliability). The one exception concerned the correlation between the BF-Agency Scale and the agency composite. That correlation dropped from $r_{da} = .81$ to $r_{da} = .77$.
7. We performed two additional analyses to probe the robustness of our results. (a) We repeated the AC method several times in order to see whether those repetitions result in the same AC-Agency-Communion Scales. (b) We doubled the number of candidate scales per iteration (60 \rightarrow 120) and the number of iterations (40 \rightarrow 80). Neither additional analysis altered the results.
8. To aid readability, we do not present p values in the main text. Instead, we note here that our smallest sample (Exploratory Sample) had a size of $n = 308$. With such a sample size and a power of 80% correlations of $.16 \leq r < .19$ are significant at the .05-level, correlations of $.19 \leq r < .23$ are significant at the .01-level, and correlations of $r \geq .23$ are significant at the .001-level. Note that Figures 1 depicts the main-text results, including 95% confidence intervals.

9. To control for acquiescence, we used the procedure suggested by Soto et al. (2011).
10. We formed three parcels for each ER-Agency-Communion Scale, following the “domain representative approach” (see Little et al., 2002, pp. 167-168).
11. Convergence between the four BFI-Agency-Communion Scales would be unremarkable, if item-overlap between the four scales were very high. Yet we did not find very high item-overlap. More precisely, the average item-overlap between the four BFI-Agency scales was 26.5%, ranging from 0% (between the ER- and BF-Agency Scales) to 75% (between the ER- and TS-Agency Scales). The average item-overlap between the four BFI-Communion scales was 60.7%, ranging from 47% (between the ER- and TS-Communion Scales) to 75% (between the TS- and BF-Communion Scales).
12. Soto and John (2017) acknowledged and addressed this issue by adding an agentic facet (assertiveness) to the BFI-2’s Extraversion Scale.
13. According to the narrower view on agency and communion (described in Footnote 1), extraversion (agreeableness) and agency (communion) are mere rotations of the same dimensions (McCrae & Costa, 1989; Trapnell & Wiggins, 1990).

References

- Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y.P. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness—Communion with warmth and morality. *Frontiers in Psychology, 7*, 1-17. <https://doi.org/10.3389/fpsyg.2016.01810>
- Abele, A. E., Uchrowski, M., Suitner, C., & Wojciszke, B. (2008). Towards an operationalization of the fundamental dimensions of agency and communion: Trait content ratings in five countries considering valence and frequency of word occurrence. *European Journal of Social Psychology, 38*(7), 1202-1217. <https://doi.org/10.1002/ejsp.575>
- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*(5), 751-763. <https://doi.org/10.1037/0022-3514.93.5.751>
- Abele, A. E., & Wojciszke, B. (2018). *Agency and communion in social psychology*. Routledge.
- Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the inventory of interpersonal problems. *Journal of Personality Assessment, 55*(3-4), 521-536. <https://doi.org/10.1080/00223891.1990.9674088>
- Anderson, C. A., Allen, J. J., Plante, C., Quigley-McBride, A., Lovett, A., & Rökkum, J. N. (2018). The MTurkification of social and personality psychology. *Personality and Social Psychology Bulletin, 45*(6), 842-850. <https://doi.org/10.1177/0146167218798821>
- Anusic, I., Schimmack, U., Pinkus, R. T., & Lockwood, P. (2009). The nature and structure of correlations among Big Five ratings: The halo-alpha-beta model. *Journal of Personality and Social Psychology, 97*(6), 1142-1156. <https://doi.org/10.1037/a0017159>
- Armstrong, R. D., Jones, D. H., & Wu, L. (1992). An automated test development of parallel tests from a seed test. *Psychometrika, 57*(2), 271-288. <https://doi.org/10.1007/BF02294509>
- Bakan, D. (1966). *The duality of human existence: Isolation and communion in Western man*. Beacon Press.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*(4), 396-403. <https://doi.org/10.1111/j.1745-6916.2007.00051.x>
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*(2), 155-162. <https://doi.org/10.1037/h0036215>
- Bi, C., Ybarra, O., & Zhao, Y. (2013). Accentuating your masculine side: Agentic traits generally dominate self-evaluation, even in China. *Social Psychology, 44*(2), 103-108. <https://doi.org/10.1027/1864-9335/a000144>
- Brewer, M. B. (2013). 25 years toward a multilevel science. *Perspectives on Psychological Science, 8*, 554-555.
- Buchanan, K., & Bardi, A. (2015). The roles of values, behavior, and value-behavior fit in the relation of agency and communion to well-being. *Journal of Personality, 83*(3), 320-333. <https://doi.org/10.1111/jopy.12106>
- Campbell, W. K., Rudich, E. A., & Sedikides, C. (2002). Narcissism, self-esteem, and the positivity of self-views: Two portraits of self-love. *Personality and Social Psychology Bulletin, 28*(3), 358-368. <https://doi.org/10.1177/0146167202286007>
- Combs, J. G. (2010). Big samples and small effects: Let’s not trade relevance and rigor for power. *Academy of Management Journal, 53*(1), 9-13. <https://doi.org/10.5465/amj.2010.48036305>
- Costa, P. T., & McCrae, R. R. (2008). The revised NEO Personality Inventory (NEO-PI-R). In G. Matthews, D. Saklofske, & G. J. Boyle (Eds.), *The SAGE handbook of personality theory and assessment: Personality measurement and testing* (Vol. 2, pp. 179-198). Sage.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. <https://doi.org/10.1007/BF02310555>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. <https://doi.org/10.1037/h0040957>
- DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91*(6), 1138-1151. <https://doi.org/10.1037/0022-3514.91.6.1138>
- DeYoung, C. G., Weisberg, Y. J., Quilty, L. C., & Peterson, J. B. (2013). Unifying the aspects of the Big Five, the interpersonal circumplex, and trait affiliation. *Journal of Personality, 81*(5), 465-475. <https://doi.org/10.1111/jopy.12020>
- 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>
- Dorigo, M., & Gambardella, L. M. (1997). Ant colonies for the travelling salesman problem. *Biosystems, 43*(2), 73-81. [https://doi.org/10.1016/S0303-2647\(97\)01708-5](https://doi.org/10.1016/S0303-2647(97)01708-5)
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Frimer, J. A., Walker, L. J., Dunlop, W. L., Lee, B., & Riches, A. (2011). The integration of agency and communion in moral

- personality: Evidence of enlightened self-interest. *Journal of Personality and Social Psychology*, 101(1), 149-163. <https://doi.org/10.1037/a0023780>
- Fournier, M. A., Di Domenico, S. I., Weststrate, N. M., Quitasol, M. N., & Dong, M. (2015). Toward a unified science of personality coherence. *Canadian Psychology/Psychologie Canadienne*, 56, 253-262.
- Funder, D. (2020). *Misgivings: Some thoughts about "measurement invariance."* <https://funderstorms.wordpress.com/2020/01/31/misgivings-some-thoughts-about-measurement-invariance/>
- Furr, R. M., & Funder, D. C. (1998). A multimodal analysis of personal negativity. *Journal of Personality and Social Psychology*, 74(6), 1580-1591. <https://doi.org/10.1037/0022-3514.74.6.1580>
- Gebauer, J. E. (2015). *Exploratory parallel parcel analysis*. Paper presentation at the expert meeting "Advances in Scale Construction in the Social Sciences" at GESIS Leibniz-Institute for the Social Sciences in Mannheim, Germany.
- Gebauer, J. E., Eck, J., Entringer, T. M., Bleidorn, W., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2020). The well-being benefits of person-culture match are contingent on basic personality traits. *Psychological Science*, 31(10), 1283-1293. <https://doi.org/10.1177/0956797620951115>
- Gebauer, J. E., Paulhus, D. L., & Neberich, W. (2013). Big Two personality and religiosity across cultures: Communals as religious conformists and agentics as religious contrarians. *Social Psychological and Personality Science*, 4(1), 21-30. <https://doi.org/10.1177/1948550612442553>
- Gebauer, J. E., Sedikides, C., Lütke, O., & Neberich, W. (2014). Agency, communion and interest in prosocial behavior: Social motives for assimilation and contrast explain sociocultural inconsistencies. *Journal of Personality*, 82(5), 452-466. <https://doi.org/10.1111/jopy.12076>
- Gebauer, J. E., Sedikides, C., Verplanken, B., & Maio, G. R. (2012). Communal narcissism. *Journal of Personality and Social Psychology*, 103(5), 854-878. <https://doi.org/10.1037/a0029629>
- Gebauer, J. E., Sedikides, C., Wagner, J., Bleidorn, W., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2015). Cultural norm fulfillment, interpersonal belonging, or getting ahead? A large-scale cross-cultural test of three perspectives on the function of self-esteem. *Journal of Personality and Social Psychology*, 109(3), 526-548. <https://doi.org/10.1037/pspp0000052>
- Gebauer, J. E., Wagner, J., Sedikides, C., & Neberich, W. (2013). Agency-communion and self-esteem relations are moderated by culture, religiosity, age, and sex: Evidence for the "self-centrality breeds self-enhancement" principle. *Journal of Personality*, 81(3), 261-275. <https://doi.org/10.1111/j.1467-6494.2012.00807.x>
- 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. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7-28). Tilburg University Press.
- 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(1), 84-96. <https://doi.org/10.1016/j.jrp.2005.08.007>
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, 59(2), 93-104. <https://doi.org/10.1037/0003-066X.59.2.93>
- Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504-528.
- Gray, K. (2017). How to map theory: Reliable methods are fruitless without rigorous theory. *Perspectives on Psychological Science*, 12(5), 731-741. <https://doi.org/10.1177/1745691617691949>
- Gurven, M., von Rueden, C., Massenkoff, M., Kaplan, H., & Lero Vie, M. (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>
- Gurven, M., von Rueden, C., Stieglitz, J., Kaplan, H., & Rodriguez, D. E. (2014). The evolutionary fitness of personality traits in a small-scale subsistence society. *Evolution and Human Behavior*, 35(1), 17-25. <https://doi.org/10.1016/j.evolhumbehav.2013.09.002>
- Guttman, L. A. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10(4), 255-282. <https://doi.org/10.1007/BF02288892>
- Hase, H. D., & Goldberg, L. R. (1967). Comparative validity of different strategies of constructing personality inventory scales. *Psychological Bulletin*, 67(4), 231-248. <https://doi.org/10.1037/h0024421>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 62-135. <https://doi.org/10.1017/S0140525X0999152X>
- Hogan, R. (1982). A socioanalytic theory of personality. In M. Page (Ed.), *Nebraska symposium on motivation* (Vol. 30, pp. 55-89). University of Nebraska Press.
- Hopwood, C. J., Wright, A. G., & Donnellan, M. B. (2011). Evaluating the evidence for the general factor of personality across multiple inventories. *Journal of Research in Personality*, 45(5), 468-478. <https://doi.org/10.1016/j.jrp.2011.06.002>
- Hunt, T. D., & Bentler, P. M. (2015). Quantile lower bounds to reliability based on locally optimal splits. *Psychometrika*, 80(1), 182-195. <https://doi.org/10.1007/s11336-013-9393-6>
- Imhoff, R., & Koch, A. (2017). How orthogonal are the Big Two of social perception? On the curvilinear relation between agency and communion. *Perspectives on Psychological Science*, 12(1), 122-137. <https://doi.org/10.1177/1745691616657334>
- 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* (Vol. 2, pp. 102-138). Guilford Press.

- Johnson, J. A. (2014). Measuring thirty facets of the five factor model with a 120-Item Public Domain Inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality, 51*(August), 78-89. <https://doi.org/10.1016/j.jrp.2014.05.003>
- Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology, 89*, 899-913.
- Krizan, Z., & Herlache, A. D. (2018). The narcissism spectrum model: A synthetic view of narcissistic personality. *Personality and Social Psychology Review, 22*(1), 3-31. <https://doi.org/10.1177/1088868316685018>
- Leary, T. (1957). *Interpersonal diagnosis of personality*. Ronald Press.
- Leite, W. L., Huang, I.-C., & Marcoulides, G. A. (2008). Item selection for the development of short forms of scales using an ant colony optimization algorithm. *Multivariate Behavioral Research, 43*(3), 411-431. <https://doi.org/10.1080/00273170802285743>
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 151-173. https://doi.org/10.1207/S15328007SEM0902_1
- Locke, K. D. (2000). Circumplex scales of interpersonal values: reliability, validity, and applicability to interpersonal problems and personality disorders. *Journal of Personality Assessment, 75*(2), 249-267. https://doi.org/10.1207/S15327752JPA7502_6
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*(3), 635-694. <https://doi.org/10.2466/pr0.1957.3.3.635>
- Louie, J. F., Kurtz, J. E., & Markey, P. M. (2018). Evaluating circumplex structure in the interpersonal scales for the NEO-PI-3. *Assessment, 25*(5), 589-595. <https://doi.org/10.1177/1073191116665697>
- 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*(3), 401-412. <https://doi.org/10.1037/0021-843X.110.3.401>
- Maio, G. R. (2010). Mental representations of social values. *Advances in Experimental Social Psychology, 42*, 1-43. [https://doi.org/10.1016/S0065-2601\(10\)42001-8](https://doi.org/10.1016/S0065-2601(10)42001-8)
- Markey, P. M., & Markey, C. N. (2009). A brief assessment of the interpersonal circumplex: The IPIP-IPC. *Assessment, 16*(4), 352-361. <https://doi.org/10.1177/1073191109340382>
- Martin, A. E., & Slepian, M. L. (2020). The primacy of gender: Gendered cognition underlies the Big Two dimensions of social cognition. *Perspectives on Psychological Science*. Advance online publication. <http://dx.doi.org/10.1177/1745691620904961>
- McAdams, D. P., Hoffman, B. J., Mansfield, E. D., & Day, R. (1996). Themes of agency and communion in significant autobiographical scenes. *Journal of Personality, 64*(2), 339-378. <https://doi.org/10.1111/j.1467-6494.1996.tb00514.x>
- McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review, 19*(2), 97-112. <https://doi.org/10.1177/1088868314541857>
- McCrae, R. R., & Costa, P. T. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology, 56*(4), 586-595. <https://doi.org/10.1037/0022-3514.56.4.586>
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175-215. <https://doi.org/10.1111/j.1467-6494.1992.tb00970.x>
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437-455. <https://doi.org/10.1037/a0028085>
- Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality, 41*(6), 1213-1233. <https://doi.org/10.1016/j.jrp.2007.02.003>
- Olaru, G., Witthöft, M., & Wilhelm, O. (2015). Methods matter: Testing competing models for designing short-scale Big-Five assessments. *Journal of Research in Personality, 59*(December), 56-68. <https://doi.org/10.1016/j.jrp.2015.09.001>
- Paulhus, D. L. (2018). Agency and communion: Two dimensions of desirability. In A. E. Abele & B. Wojciszke (Eds.), *Agency and communion in social psychology* (pp.123-129). Routledge.
- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality, 66*(6), 1025-1060. <https://doi.org/10.1111/1467-6494.00041>
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224-239). Guilford Press.
- Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods, 46*(4), 1023-1031. <https://doi.org/10.3758/s13428-013-0434-y>
- Pettersson, E., Turkheimer, E., Horn, E. E., & Menatti, A. R. (2012). The general factor of personality and evaluation. *European Journal of Personality, 26*(3), 292-302. <https://doi.org/10.1002/per.839>
- Rau, R., Carlson, E. N., Back, M. D., Barranti, M., Gebauer, J. E., Human, L. J., Leising, D., & Nestler, S. (2019, December 12). What is the structure of perceiver effects? On the importance of global positivity and trait-specificity across personality domains and judgment contexts. *Journal of Personality and Social Psychology*. Advance online publication. <http://dx.doi.org/10.1037/pspp0000278>
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.
- Russell, S. S., Spitzmüller, C., Lin, L. F., Stanton, J. M., Smith, P. C., & Ironson, G. H. (2004). Shorter can also be better: The abridged job in general scale. *Educational and Psychological Measurement, 64*(5), 878-893. <https://doi.org/10.1177/0013164404264841>
- Saucier, G., Thalmayer, A. G., & Bel-Bahar, T. S. (2014). Human attribute concepts: Relative ubiquity across twelve mutually isolated languages. *Journal of Personality and Social Psychology, 106*(2), 287-300. <https://doi.org/10.1037/a0035888>

- Psychology*, 107(1), 199-216. <https://doi.org/10.1037/a0036492>
- Saucier, G., Thalmayer, A. G., Payne, D. L., Carlson, R., Sanogo, L., Ole-Kotikash, L., A. Church, T., Katigbak, M. S., Somer, O., Szarota, P., Zsofia 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>
- Saucier, G. (1994). Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment*, 63, 506-516.
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609-612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Şimşek, Ö. F. (2012). Higher-order factors of personality in self-report data: Self-esteem really matters. *Personality and Individual Differences*, 53(5), 568-573. <https://doi.org/10.1016/j.paid.2012.04.023>
- 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(1), 117-143. <https://doi.org/10.1037/pspp0000096>
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, 100, 330-348.
- Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates, and antecedents*. University of Texas Press.
- Srivastava, S., Guglielmo, S., & Beer, J. S. (2010). Perceiving others' personalities: Examining the dimensionality, assumed similarity to the self, and stability of perceiver effects. *Journal of Personality and Social Psychology*, 98(3), 520-534. <https://doi.org/10.1037/a0017057>
- Stanton, J. M. (2000). Empirical distributions of correlations as a tool for scale reduction. *Behavior Research Methods, Instruments, & Computers*, 32(3), 403-406. <https://doi.org/10.3758/BF03200808>
- Stanton, J. M., Sinar, E. F., Balzer, W. K., & Smith, P. C. (2002). Issues and strategies for reducing the length of self-report scales. *Personnel Psychology*, 55, 167-194.
- Suitner, C., & Maass, A. (2008). The role of valence in the perception of agency and communion. *European Journal of Social Psychology*, 38(7), 1073-1082. <https://doi.org/10.1002/ejsp.525>
- Thalmayer, A. G., Saucier, G., Ole-Kotikash, L., & Payne, D. (2020). Personality structure in east and west Africa: Lexical studies of personality in Maa and Supyire-Senufo. *Journal of Personality and Social Psychology*, 119(5), 1132-1152. <https://doi.org/10.1037/pspp0000264>
- Trapnell, P. D., & Paulhus, D. L. (2012). Agentic and communal values: Their scope and measurement. *Journal of Personality Assessment*, 94(1), 39-52. <https://doi.org/10.1080/00223891.2011.627968>
- Traubman, E. K., Smith, T. W., Uchino, B. N., Berg, C. A., Trobst, K. K., & Costa, P. T. (2009). Interpersonal circumplex octant, control, and affiliation scales for the NEO-PI-R. *Personality and Individual Differences*, 47(5), 457-463. <https://doi.org/10.1016/j.paid.2009.04.018>
- Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles*, 36(5-6), 305-325. <https://doi.org/10.1007/BF02766650>
- Van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, 44(3), 315-327. <https://doi.org/10.1016/j.jrp.2010.03.003>
- von Rueden, C. R., Lukaszewski, A. W., & Gurven, M. (2015). Adaptive personality calibration in a human society: Effects of embodied capital on prosocial traits. *Behavioral Ecology*, 26(4), 1071-1082. <https://doi.org/10.1093/beheco/arv051>
- Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In D. Cicchetti & W.M. Grove (Eds.), *Thinking clearly about psychology: Essays in honor of Paul E. Meehl, Vol.1.Matters of public interest; Vol.2.Personality and psychopathology* (pp. 89-113). University of Minnesota Press.
- Wiggins, J. S. (2003). *Paradigms of personality assessment*. Guilford Press.
- Wiggins, J. S., & Trapnell, P. D. (1996). A dyadic-interactional perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 88-93). Guilford Press.
- Wiggins, J. S., Trapnell, P., & Phillips, N. (1988). Psychometric and geometric characteristics of the Revised Interpersonal Adjective Scales (IAS-R). *Multivariate Behavioral Research*, 23(4), 517-530. https://doi.org/10.1207/s15327906mbr2304_8
- Woods, S. A., & Hampson, S. E. (2005). Measuring the Big Five with single items using a bipolar response scale. *European Journal of Personality*, 19(5), 373-390. <https://doi.org/10.1002/per.542>
- Wojciszke, B., Baryła, W., Parzuchowski, M., Szymkow, A., & Abele, A. E. (2011). Self-esteem is dominated by agentic over communal information. *European Journal of Social Psychology*, 41(5), 617-627. <https://doi.org/10.1002/ejsp.791>
- Yarkoni, T. (2010). The abbreviation of personality, or how to measure 200 personality scales with 200 items. *Journal of Research in Personality*, 44(2), 180-198. <https://doi.org/10.1016/j.jrp.2010.01.002>