Supplemental Material S1

Table S1.
Descriptions of the studies included in the meta-analysis.

Author	Relationship Type	*N* targets
Carroll (1952)	cohabitators	110
McEwan & Devins (1983)	friends	7-19
Malloy & Albright (1990)	cohabitators	84
Watson & Clark (1991)	other	150
Davidson (1993)	friends	92
Mount, Barrick, & Strauss (1994)	work colleagues	92-105
White & Nias (1994)	family	110
Botwin, Buss, & Shackelford (1997)	family, friends, strangers	105-106
Foltz, Morse, Calvo, & Barber (1997)	family	98
Gomà-i-Freixanet (1997)	family	198
Kosek (1998)	family	108
Spirrison & Choi (1998)	friends	45
Chambers (1999)	family, friends, strangers	108
Barrick, Patton, & Haugland (2000)	friends, strangers	73
Lanthier (2000)	family	240
Watson, Hubbard, & Wiese (2000)	family, friends	148-558
Lang, Lüdtke, & Asendorpf (2001)	friends	126
O'Connor, Archer, & Wu (2001)	family	77
Rovner & Casten (2001)	family	10-26
Trierweiler, Eid, & Lischetzke (2002)	friends*	482
Zarevski, Bratko, Butković, & Lazić (2002)	friends	111-225
Collani & Herzberg (2003)	friends	127
Hendriks et al. (2003)	friends	97

Kurtz & Sherker (2003)	friends	216
Lee, Gizzarone, & Ashton (2003)	friends	150
Saroglou & Fiasse (2003)	family	122
Besser (2004)	friends	187
Biesanz & West (2004)	family, friends	256
Gomà-i-Freixanet, Wismeijer, & Valero (2005)	family	171
Longley, Watson, & Noyes (2005)	family	171
Bratko, Chamorro-Premuzic, & Saks (2006)	friends	255
Davis & Mattoon (2006)	friends	187
Holden, Wasylkiw, Starzyk et al. (2006)	cohabitators	420
Lischetzke & Eid (2006)	friends	475
Small & Diefendorff (2006)	work colleagues	143
Watson & Humrichouse (2006)	family	306-574
Biesanz, West, & Millevoi (2007)	friends	184
Furr, Dougherty, Marsh, & Mathias (2007)	family	53
Lönnqvist, Paunonen, Tuulio-Henriksson et al. (2007)	family	70
Muck, Hell, & Gosling (2007)	friends	175
Wagerman & Funder (2007)	friends	10-57
Yeagley, Morling, & Nelson (2007)	strangers	6
Back, Schmukle, & Egloff (2008)	strangers	599
Holland & Roisman (2008)	friends	80-218
Kurtz, Tarquini, & Iobst (2008)	family, friends	89
Tahmasb, Ghorbani, & Watson (2008)	friends	288
Vazire, Naumann, Rentfrow, & Gosling (2008)	family, friends*	41-160
Borkenau, Zaltauskas, & Leising (2009)	friends	284
Jackson, Bogg, Walton et al. (2009)	other	74-124
Soto & John (2009)	friends	590

Back, Schmukle, & Egloff (2010)	strangers	72-73
Back, Stopfer, Vazire et al. (2010)	strangers	103
Borkenau, Paelecke, & Yu (2010)	other	126
Gros, Simms, & Antony (2010)	friends	127
Joseph & Newman (2010)	friends	280
Kandler, Riemann, Spinath, & Angleitner (2010)	family, friends	607
Leising, Erbs, & Fritz (2010)	friends	90
Letzring & Noftle (2010)	family, friends	268
Simms, Zelazny, Yam, & Gros (2010)	friends	606
Vazire (2010)	friends, strangers	165
Voracek, Gabler, Kreutzer et al. (2010)	friends	48
Bleidorn & Peters (2011)	work colleagues	44
Carlson, Vazire, & Furr (2011)	friends, strangers*	68-198
Leising (2011)	strangers	108
Marini & Kurtz (2011)	family, friends	103-128
Miller, Jones, & Lynam (2011)	friends	64
Traupman, Smith, Florsheim et al. (2011)	family	300
Colbert, Judge, Choi, Wang (2012)	friends	155
Connelly & Hülsheger (2012)	friends, work colleagues	97
Connelly & Wollscheid (2012)	friends	101
Hong, Koh, & Paunonen (2012)	friends	246
Kim & Schimmack (2012)	friends	922-947
Koydemir & Schütz (2012)	friends	86-101
Kurtz, Puher, & Cross (2012)	family, friends	106
Nauta (2012)	friends	114
Nestler, Egloff, Küfner, & Back (2012)	strangers	18-40
Beer, Watson, & McDade-Montez (2013)	family, friends	124-381

Buschor, Proyer, Ruch (2013)	friends	330
Cohen, Panter, Turan et al. (2013)	friends, work colleagues*	320-341
Di Domenico, Quitasol, & Fournier (2013)	strangers	249
Fleischhauer, Enge, Miller et al. (2013)	friends	204
Gallrein, Carlson, Holstein, & Leising (2013)	friends	65
Hall & Pennington (2013)	strangers	100
Klafehn, Li, & Chiu (2013)	friends*	206
Leikas, Lönnqvist, Verkasalo, Nissinen (2013)	work colleagues	35
Leising, Borkenau, Zimmermann et al. (2013)	other	129
Momm, Blickle, & Liu (2013)	work colleagues	123
Orth (2013)	cohabitators	186
Siers & Christiansen (2013)	friends	200
Sorokowska (2013)	strangers	50
Webb, Nelson, Huelsman, Bleske-Rechek (2013)	friends	198
Clifton (2014)	family, friends	52-111
de Vries, Zettler, & Hilbig (2014)	family, friends	1106
Saeki, Oishi, Maeno, & Gilbert (2014)	friends	202
Uziel (2014)	family	70
Dufner, Egloff, Hausmann et al. (2015)	friends*	106
Dermody, Wright, Cheong et al. (2016)	family, friends*	856
Kholin, Meurs, Blickle et al. (2016)	friends	116
Solomon & Vazire (2016)	friends	173
Allik data	family	2679-2733
Ashton data	friends	2137
Beer data	friends, strangers*	76-107
Eugene-Springfield Community Sample	family, friends, work colleagues, other	41-656
DeYoung data	friends	131
Furr data	family, friends, strangers*	13-291

Letzring data	other	56
Page-Gould data	strangers	100-241
Petrican & Schimmack	friends	214
Powell data	friends	254
Schimmack data	family	404-410
Vazire data	family, friends*	13-217
Walker & Schimmack data	friends	127-310
Zou & Schimmack data	family, friends, other	131-191

Note. *Informants can include other types of acquaintances.

Table S2.
Hough and Ones' (2001) Big Five taxonomy with example inventories.

Emotional Stability	Big Five Inventory: Neuroticism
Anxiety	PANAS-X: Fear
Uneven-temperedness	NEO-PI-R: Anger
Skepticism	NEO-3: Angry Hostility
Depression	Guilt and Shame Proneness scale: Shame Proneness
Self-Esteem	Rosenberg's Self-Esteem Scale
Negative Affect	Temperament and Emotion Questionnaire: Negative Affect
Extraversion	HEXACO: Extraversion
Sensation Seeking	NEO-3: Excitement Seeking
Sociability	Self-Directed Search: Social Introversion (Holland, Fritzsche, & Powell, 1997)
Positive Emotions	Positive and Negative Affect Scales: Positive Affect
Dominance	Interpersonal Adjective Scale: Assured-Dominant
Activity	Zuckerman-Kuhlman Personality Questionnaire: Activity
Openness/Intellect	Trait Adjectives Extracted from Saucier and Ostendorf (1999; Table 2)
Curiosity	Big Five Inventory: Aesthetics
Depth	NEO-3: Fantasy
Aesthetics	Values in Action Inventory of Strengths: Beauty
Openness to Sensations	Bipolar Adjective Rating Scale: Novelty/Openness (Johnson, 2000)
Intellectual Efficiency	One item selected from Norman's (1963) five higher order factors: Intellectual
Nontraditionalism	Supernumerary personality Inventory: Conventionality (Paunonen, 2002)
Agreeableness	Mini markers: Agreeableness
Nurturance	Interpersonal Adjectives Scale Revised: Cold-Hearted
Non-manipulative	One item measure: Is an honest soul (Gallrein, Carlson, Holstein, & Leising, 2013)
Cooperation	Big Five Inventory: Compliance

Modesty	Values in Action Inventory of Strengths: Modesty
Conscientiousness	NEO-FFI: Conscientiousness
Achievement	Conscientiousness Adjective Checklist: Industriousness (Roberts, O'Donnell, & Robins, 2004)
Cautiousness	Self-Directed Search: Deliberation (Holland et al., 1997)
Dependability	Big Five Inventory: Self-Discipline
Orderliness	Behavioral Indicators of Conscientiousness : Orderliness (Jackson et al. 2010)

Note. The current taxonomy is based on Hough and Ones' (2011) original taxonomy that has been used in many meta-analytic reviews to categorize a breadth of personality scales as measures of particular Big Five factors, facets, or interstitial/compound traits (e.g., Dudley, Orvis, Lebiecki, & Cortina, 2006; Foldes, Duehr, & Ones, 2008; Trapmann, Hell, Hirn, & Schuler, 2007). For example, the NEO-PI-R Extraversion domain is classified to *Global Extraversion* taxon, whereas the NEO-PI-R Gregariousness facet is classified to *Extraversion: Sociability* taxon. The present meta-analysis used existing and emerging large-scale multi-inventory research to more cleanly distinguish facet traits and to delineate them from compound/interstitial traits.

Supplemental Material S2

Collection method

The current meta-analysis used a range of different methods to collect the data. First, we

reviewed 431 articles compiled in Connelly and Ones' (2010) meta-analytic database of

informant inter-rater reliabilities, self-informant correlations, and informant correlations with

behaviors. Second, we updated this database by searching in *PsycInfo* to locate articles that had

been published since 2008. In this *PsycInfo* search, we used the following search string:

(personality or trait or temperament) and (peer or informant or spouse or friend or roommate or

stranger or consensual validity or consensual validation or self-informant agreement or zero

acquaintance or thin slices of behavior). The search results were assessed by the second author

for eligibility by scanning the titles and abstracts. The detailed search produced 297 articles for

potential review. Third, we reviewed the reference sections of additional meta-analyses of

informant-report studies (Kenny, Albright, Malloy, & Kashy, 1994; Oh, Wang, & Mount, 2011).

Fourth, we examined psychological test manuals for potential comparisons between self- and

informant-report forms (including the California Personality Inventory (3rd Edition), HEXACO

Manual, HPI Manual, JPI Jackson Personality Inventory manual, OPQ32 Technical Manual, and

NEO Inventories – NEO-PI-3, NEO-FFI-3, and NEO-PI-R). Fifth, we reviewed the abstracts of

relevant conference programs since 2008 (Association for Research in Personality, European

Association of Personality Psychology, International Society for the Study of Individual

Differences, and Society for Personality and Social Psychology). Sixth, we contacted 33

researchers who have frequently published work using self and informants to request

unpublished data. Lastly, we posted a call for working papers, forthcoming articles, and

unpublished data on relevant list serves and associations' websites (the Society for Personality

and Social Psychology, Association for Research in Personality, Society of Australasian Social

Psychologists, Canadian Psychological Association, Society for the Psychological Study for

Social Issues, and the Society for Industrial and Organizational Psychology).

In many cases, the articles and manuscripts we obtained did not report specific statistics

of interest (i.e., means and standard deviations) but appeared to have collected self- and

informant-report data. In such cases, we requested further information from the corresponding

authors of any articles published since 2000.

Inclusion and exclusion criteria

There were six inclusion/exclusion criteria to select the studies for the meta-analysis.

First, the study must have collected personality reports from both self and an informant source

using the same scale and reported their means, standard deviations, and sample size. For

example, studies in which the self completed a personality inventory on likert-type scale and

informant(s) completed the same inventory on adjective-anchored rating scale were excluded.

Second, raters must be rating the same targets. We excluded studies that asked half of the sample

to rate themselves and the other half to rate someone they know well as these two reports are not

paired and their discrepancy scores reflect assumed similarity (lack objective criteria). Third, we

included only articles reporting data on "normal" range personality traits (i.e., scales designed to

measure psychopathology were excluded). This facilitated our focus on normal personality traits

that tend to be normally distributed in the general populations, and align with the Big Five

personality studies. Fourth, we excluded traumatized or heavy clinical samples to avoid

populations whose self- or informant-perspectives may be particularly distorted. Fifth, we also

excluded studies using ipsative measures and/or peer-nomination procedures in which a group of

raters (e.g., a fraternity) nominated the group member who scored highest or lowest on a

particular trait. These measures entail psychometric problems caused by an artificial dependence among the scales (e.g., resulting in inflation of reliability coefficients). Lastly, we only included studies that provided the overall sample sizes, means, standard deviations of the self and informant source to correct for unreliability. If a sample was used in more than one article, we coded it once to ensure independence of effect sizes.

Database Coding

Personality Variables. First, we coded the inventory according to the working taxonomy of Big Five factors and facets. All authors coded all personality scales according to this taxonomy, classifying scales both within a particular Big Five domain and as a global measure or particular facet/compound trait. Inter-rater reliabilities between the coders were strong, with Kappa κ=.914 and κ=.892 (Fleiss, 1971; Geertzen, 2012). In cases of disagreements between three authors, final categorization was determined by discussion between all three coders until agreement was reached.

Study Characteristics. In addition to the personality measure, we identified characteristics of the sample and type of relationship that informants had with the target. To assess inter-rater reliability in the coding, two raters coded a subset of studies (10% of studies reviewed). If the studies did not report the necessary information, we contacted the corresponding author for additional information.

We conducted inter-rater reliability assessments for key data elements using coefficient Kappa κ (Cohen, 1960) for categorical variables and intraclass correlation (ICC) for numerical variables (to examine the agreement between the two judges). Inter-rater agreement was good for Big Five classification (κ=1), and informant type (κ=.80), and for continuous variables ranging

from .9 to 1 (e.g., means and standard deviations). Any discrepancies between raters were resolved through discussion.

Formula. Formula to correct for number of raters for averaged scores and summed scores of the standard deviation and reliability scores. $S_x^2 =$ variance of a single rater, k =number of raters per target, $S_{\bar{x}}^2$=variance of the average across k raters, and r_{xx} = interrater reliability of a single rater.

For averaged scores:
$$\sigma_x^2 = \frac{k\sigma_{\bar{x}}^2}{1 + r_{xx}(k-1)}$$

For summed scores:
$$\sigma_x^2 = \frac{\sigma_{\Sigma x}^2}{k[r_{xx}(k-1)+1]}$$

Standard deviations of self- and informant-reports were disattenuated before they were put into the two formulas of interest if any of the estimates were based on more than one rater. To correct for the bias, we used inter-rater reliability scores reported in the article. If the study reported reliability scores for multiple raters, reliability scores were attenuated for number of raters (r_{xx}=((1/k)* $\bar{r}$)/(1+((1/k)-1)* $\bar{r}$; $\bar{r}$ =interrater reliability of multiple raters). If interrater reliability were not available in the article, we used the interrater reliability estimates reported in Connelly's (2008) dissertation.

Winsorizing. In a meta-analytic review, the results of extremely larger studies can disproportionately affect the results. Thus, we winsorized the three large sample size studies to prevent contributions of larger studies affecting the results. There were three outlying studies that

were winsorized to 2.5 SDs ($N = 1,006$) from the mean of all sample sizes. Identical moderator

analyses were conducted without changing sample sizes of the three studies and produced similar

results.

Results

      Publication Bias. Publication bias was calculated using the Egger′s test (Egger, Smith,

Schneider, & Minder, 1997) and Begg-Mazumdar test (Begg & Mazumdar, 1994) to detect

detecting asymmetry in the funnel plots. All analyses were carried out in the R environment for

statistical computing and visualization. Roughly symmetric funnel plots would suggest low risk

of publication bias. Results suggest no evidence of publication bias. The null hypothesis for

Egger′s test is that symmetry is present with the alternative hypothesis indicating that asymmetry

is present. All p-values were above .05. Thus, there was no evidence to reject the null hypothesis

in favor of the alternative suggesting that there was no publications bias in the studies included in

the meta-analysis. The Begg-Mazumdar test confirmed these observations with non-significant

results. This method is used to calculate a rank-correlation between effect size and its

corresponding standard error. A significant and strong correlation indicates presence of

publication bias (e.g., larger effects are disproportionately represented in smaller sample studies),

and all Begg−Mazumdar indicators gave Kendall's tau values close to zero suggesting no

publication bias.

Table S3.
Publication test results.

	Egger's regression test		Begg-Mazumdar correlation test	
	z	*p*	Kendall's τ	*p*
Emotional Stability	-.450	.653	-.003	.950
Extraversion	-.068	.946	.027	.518
Openness/Intellect	.282	.778	.059	.214
Agreeableness	-.277	.782	-.022	.646
Conscientiousness	-.656	.512	-.057	.228

Supplemental Material S3

Table S4.

Independent raters´ratings of level of desirability and observability of the traits and facets.

Traits	Desirability M	Desirability SD	Observability M	Observability SD
Emotional Stability	5.68	1.60	3.21	1.23
Extraversion	5.58	.84	6.47	.51
Openness to Experience	5.26	.81	3.84	1.30
Agreeableness	6.05	.78	4.32	1.16
Conscientiousness	6.42	.61	5.21	.79
Facets				
Anxiety	1.95	.41	3.53	1.39
Uneven-temperedness	2.37	1.17	4.47	1.12
Skepticism	4.00	1.05	3.68	1.46
Depression	1.63	.60	3.21	1.27
Self-esteem	6.11	.74	3.95	1.18
Negative Affect	2.00	.75	4.11	1.52
Sensation Seeking	4.05	.41	5.05	1.22
Sociability	6.16	.60	6.37	.60
Positive Emotions	6.58	.61	5.68	.48
Dominance	4.68	.95	6.16	.60
Activity	5.16	.83	6.26	1.05
Curiosity	5.74	.81	3.68	1.29
Depth	5.95	1.03	2.84	1.30
Aesthetics	5.05	1.03	3.47	1.47

Openness to Sensations	4.63	.90	3.11	1.41
Intellectual Efficiency	6.63	.68	3.37	1.54
Nontraditionalism	4.26	.73	4.84	1.21
Nurturance	5.63	.76	4.47	1.12
Non-manipulative	6.37	.76	3.68	1.29
Cooperation	6.26	.73	5.21	1.13
Modesty	5.37	1.01	4.68	1.11
Achievement	6.21	.71	5.16	1.30
Cautiousness	4.95	1.81	4.63	.90
Dependability	6.37	1.17	4.47	1.22
Orderliness	5.21	.63	5.95	1.18

Note. $N = 19$. *Desirability* = Independent raters' ratings of level of desirability. Observability = Independent raters' ratings of level of observability of the traits and facets. We used an adaptation of John & Robins (1993) method for scoring the evaluativeness of traits. Specifically, we calculated evaluativeness as the absolute difference between the average expert desirability rating and the midpoint of the scale. Additionally, visibility was based on the observability rating.

Supplemental Material S4

We conducted additional meta-analyses on the different types of personality measures. Overall, same patterns of mean differences were observed across personality measures and scales. The largest mean differences were observed in Openness/Intellect domain across measures except for trait adjective measures. Interestingly, self-report means were higher than informant-report means for Agreeableness and Conscientiousness among trait adjective measures (vs. short phrase measures). The sample size for trait adjective measures was rather small thus, further studies with larger sample sizes will be required to confirm the present findings. In general, there was minimal reduction in variability of effect sizes suggesting that our conclusions appear to be robust across different personality measures.

Table S5.
Meta-analyses of the Big Five traits across commonly used measures.

Measure and Trait	k	N	Mean Differences				Variability Ratios	
			d	SD_d	δ	SD_δ	U	SD_U
Trait Adjectives								
Emotional Stability	34	7,756	-.079	.254	**-.091**	.198	**.982**	.115
Extraversion	37	8,502	-.112	.271	**-.130**	.228	**1.007**	.108
Openness/Intellect	14	2,134	.004	.267	**.005**	.163	**.930**	.118
Agreeableness	17	2,613	.110	.218	**.132**	.000	**.858**	.176
Conscientiousness	17	2,168	.102	.256	**.123**	.062	**.852**	.079

Short Phrases								
Emotional Stability	105	23,831	-.130	.218	**-.150**	.127	**1.043**	.209
Extraversion	99	23,477	-.096	.177	**-.111**	.000	**1.005**	.201
Openness/Intellect	92	22,506	.222	.236	**.266**	.179	**1.022**	.189
Agreeableness	90	21,748	-.050	.244	**-.059**	.194	**.932**	.202
Conscientiousness	97	22,656	-.128	.253	**-.154**	.208	**.941**	.188
BFI scales								
Emotional Stability	42	9,955	-.151	.269	**-.174**	.227	**1.000**	.183
Extraversion	43	10,069	-.162	.191	**-.188**	.054	**1.001**	.204
Openness/Intellect	40	9,661	.133	.274	**.159**	.245	**.951**	.075
Agreeableness	42	9,853	-.056	.280	**-.067**	.251	**.901**	.189
Conscientiousness	49	9,892	-.138	.311	**-.167**	.294	**.921**	.201
NEO scales								
Emotional Stability	32	6,993	-.121	.197	**-.141**	.053	**1.079**	.215
Extraversion	31	7,200	-.035	.145	**-.041**	.000	**1.048**	.238
Openness/Intellect	30	7,067	.336	.141	**.402**	.000	**1.160**	.254
Agreeableness	28	6,531	-.020	.216	**-.023**	.132	**.977**	.260
Conscientiousness	29	6,951	-.166	.200	**-.201**	.096	**.967**	.211

Note. Trait Adjectives = 4 unipolar trait adjectives (Trierweiler, Eid, & Lischetzke, 2002), 30 unipolar trait adjectives (Ostendorf, 1990); Short Phrases= Big Five Inventory (John & Srivastava,1999); BFI scales = BFI-44, BFI-10; NEO scales = NEO FFI, NEO-PI-R, NEO-3 scales.