Why are Moderators of Self-Other Agreement Difficult to Establish?

Jüri Allik, Reinout E. de Vries, Anu Realo

PII: S0092-6566(16)30048-4
DOI: http://dx.doi.org/10.1016/j.jrp.2016.05.013
Reference: YJRPE 3553

To appear in: *Journal of Research in Personality*

Received Date: 5 February 2016
Revised Date: 15 May 2016
Accepted Date: 23 May 2016

Please cite this article as: Allik, J., de Vries, R.E., Realo, A., Why are Moderators of Self-Other Agreement Difficult to Establish?, *Journal of Research in Personality* (2016), doi: http://dx.doi.org/10.1016/j.jrp.2016.05.013

Running head: MODERATORS OF SELF-OTHER AGREEMENT

Why are Moderators of Self-Other Agreement Difficult to Establish?

Jüri Allik [1,2]

Reinout E. de Vries [3,4]

Anu Realo [1,5]

[1]Department of Psychology, University of Tartu

[2]The Estonian Academy of Sciences

[3]Department of Social and Organizational Psychology, Vrije Universiteit Amsterdam

[4]Department of Psychology, University of Twente

[5]Department of Psychology, University of Warwick

Corresponding author:

Jüri Allik

Department of Psychology

University of Tartu

Näituse 2, Tartu 50409

Estonia

E-mail: juri.allik@ut.ee

Abstract

Agreement between the self and other rated personality profiles was studied in two samples involving 11,096 speakers of two languages, Dutch and Estonian, who completed two different personality questionnaires, the *NEO-PI-3* and *HEXACO-PI-R*. An outstanding agreement was achieved in the most occasions: in only 4-6% of dyadic pairs was the correlation between two randomly paired profiles higher than the actually observed correlation between true pairs. As in previous studies, we found that age and sex of participants and length of acquaintance had no significant effect on the level of self-other agreement. However, intimate knowledge helped married and unmarried couples in both samples be more accurate in their personality judgments; family members, in turn, had knowledge that made them more accurate than two people who were just acquaintances or friends. We believe that these outcomes can be explained by the contention that the judgment of another's personality is a relatively simple task, which is accomplishable for most people most of the time. In other words, because judging another person's personality is an easy task, we are not able to determine "good targets," "good judges," or "good traits." Perhaps it is only "good information" which determines the closeness of the target-judge relationship, and which has a small but reliable impact on the level of self-other agreement. This explains why it is so difficult to find individual differences in the ability to judge another person's personality.

Keywords: Personality traits; self-other agreement; moderator variables; NEO-PI-3; HEXACO-PI-R; personality judgements

Why are Moderators of Self-Other Agreement Difficult to Establish?

Although Louis Thurstone was optimistic about the development of factorial methods of analysis that could give us the tools by which to reduce the complexities of social and psychological phenomena to a limited number of elements (Thurstone, 1934), most personality researchers have remained quite skeptical about the possibility of being able to decipher personality structure. Even Gordon Allport, the founder of modern trait psychology, believed that the structure of personality, that is, what personality psychologists have attempted to establish, is incredibly complex: "Since traits, like all intervening variables, are never directly observed but only inferred, we must expect difficulties and errors in the process of discovering their nature. The incredible complexity of the structure we seek to understand is enough to discourage the realist, and to tempt him to play some form of positivistic gamesmanship" (Allport, 1966, p. 3). It took many years before it became widely accepted that the huge personality lexicons expressed across the world's languages could be reduced to five (Goldberg, 1993) or six (Ashton et al., 2006; De Raad et al., 2014) independent dimensions. In other words, even though there are thousands of adjectives to describe personality dispositions in almost every spoken language, people use these words as if there were only five or six self-sufficient categories (De Raad et al., 2014; John, Angleitner, & Ostendorf, 1988; Lee & Ashton, 2008).

If personality structure, traditionally, looked incredibly complex from the viewpoint of researchers, then it was also not realistic to expect that a layperson could manage to solve the same task without facing serious problems. One of the leaders in the field, David Funder, expressed the opinion that "the accuracy of personality judgment is an extremely complex matter" (Funder, 1995, p. 653), meaning that accurate personality judgment is an unlikely

outcome which happens only when favorable circumstances are met. It is not necessary to assume, as Funder wrote, that the personality judgements are usually accurate, or even often accurate. All that is required is that lay perceivers have ever, even once, achieved accuracy in personality judgement (Funder, 1999, p. 119).

According to the *Realistic Accuracy Model* (RAM), to make an accurate personality judgment, four conditions need to be satisfied (Funder, 1995, 1999, 2012). First, the person being judged must do something relevant to exhibit the trait. Second, the trait-relevant behavior must be available to the judge. Third, the trait-relevant, available behavior must be registered. Fourth, the trait-relevant, available, and detected information must be utilized correctly to make accurate inferences about that trait (Funder, 2012). This means that personality judgment can be only conditional – at least four groups of moderating variables determine the degree to which personality judgments are accurate. Because the path to an accurate personality judgment is fundamentally uncertain, it is likely that success or failure in judgment also depends on some moderating variables. Yet, at a minimum, accuracy of judgment can be achieved with the co-occurrence of a particular set of favorable circumstances: a "good target," who possesses "good traits," who is observed by a "good judge," whose judgments are based on a "good information," and who thus applies the right inferences (Funder, 2012).

Unexpectedly, individual differences in judgeability – perhaps the most interesting among moderator variables – have been surprisingly difficult to establish (e.g., Funder, 1999; Haselton & Funder, 2006; D. J. Schneider, Hastorf, & Ellsworth, 1979). For example, already in 1937, Gordon Allport asked the famous question: "Who are these people?" (Allport, 1937, p. 443; Human & Biesanz, 2013), meaning people whose actions, thoughts, and feelings are easier to understand and judge than others (Human & Biesanz, 2013). More than 70 years later, Human

and Biesanz were obliged to admit that "despite some very interesting and important findings regarding judgeability, there still is not a clear answer to the question of who these people are" (Human & Biesanz, 2013, p. 248).

Similar problems were encountered in an attempt to identify who the "good judges" of other people's personality are. Although some subtle individual differences in the ability to judge others' personality have been observed, no large between-individual differences have been discovered in the judgment of personality traits (Bayne, 1985; Christiansen, Wolcott-Burnam, Janovics, Burns, & Quirk, 2005; Ickes, Buysse, et al., 2000; Letzring, 2008; McLarney-Vesotski, Bernieri, & Rempala, 2011; Taft, 1955). There are, of course, individual differences in empathic accuracy – another name for self-other agreement – but they do not seem to be either large or systematic (Davis & Kraus, 1997; Ickes, 1997). The only notable moderating factor for making judgments seems to be judges' intelligence: individuals with higher cognitive abilities tend to achieve slightly higher accuracy than those with lower cognitive abilities (Davis & Kraus, 1997; Murphy & Hall, 2011; Realo et al., 2003; Taft, 1955). In general, however, the ability to judge other people's personality seems to be very egalitarian: it matters very little how old you are, what your sex is (Ickes, Gesn, & Graham, 2000), or how intelligent you are (R. Mõttus, Allik, & Pullman, 2007). Although, for example, it was reported that people with elementary education may have a small advantage over those who have university degrees in judging personality (Kraus, Cote, & Keltner, 2010), this advantage is neither large nor convincingly explained. Thus, the identification of a distinctive group of people who are "good judges" has remained tentative at best.

Yet, the idea that some traits are easier to judge than others is one of the most stubborn ideas in personality psychology. It is a kind of dogma that visible traits, such as Extraversion, are more

readily assessed from an external viewpoint than less visible traits, such as Neuroticism (Connelly & Ones, 2010; Connolly, Kavanagh, & Viswesvaran, 2007; Kenny & West, 2010; McCrae et al., 2004). According to the self-other knowledge asymmetry (SOKA) model, for instance, the self should be more accurate than others for traits low in observability (e.g., Neuroticism), whereas others should be more accurate than the self for traits high in evaluativeness (e.g., Openness) (Vazire, 2010). Visibility of traits seems to provide a straightforward explanation why greatest interjudge agreement is typically on the traits that seem most observable, and lowest agreement is achieved on those traits that are not so directly observable from the vantage point of an external viewer (Funder & Dobroth, 1987; John & Robins, 1993). However, self-other agreement cannot be explained by the fact that people see themselves differently to how they are seen by other people: the disparity between self and external perspectives is unrelated to the visibility (or observability) of personality traits when correcting for other factors (Allik, Realo, Mõttus, Borkenau, et al., 2010; Allik, Realo, Mõttus, Esko, et al., 2010; Paunonen, 1989). That is, a far more important factor in self-other agreement is interindividual variance, not trait visibility. Approximately one-half of the variance in agreement level is explained by the standard deviation of the sum scores of the subscales: self-observer agreement is higher in the subscales on which individual differences are larger. After correction for the range of variance, differences in self-observer agreement are substantially diminished, although not entirely absent (Allik, Realo, Mõttus, Esko, et al., 2010). In other words, if we take into account the size of individual variation, then judges who know their target well reach more or less equal level of agreement on all personality traits usually studied by personality psychologists.

Although the idea that there are "good" targets, judges, and traits – clearly distinct from "bad" targets, judges, and traits – is very popular, very little solid evidence speaks in its favor. When reviewing the facts, Chaplin (1991) concluded that moderator effects in personality judgments are small and can only be detected in very large samples with predictors that are strongly related to the criteria (Chaplin, 1991). For example, in most cases, moderator variables are not able to transform a weak form of self-other agreement into a strong one (Chaplin, 1991). Based on the evidence, thus, individual differences in self-other agreement seem to be relatively small and, hence, difficult to detect. A telling example is the effect of the type and length of acquaintance. Although acquaintance length increases accuracy in personality judgment, there is evidence that familiarity may also have a negative effect on self-other agreement (Kenny & West, 2010). The effect of length of acquaintance is neither very consistent nor large (Bernieri, Zuckerman, Koestner, & Rosenthal, 1994; Biesanz, West, & Millevoi, 2007; Gnambs, 2013; Kenny, 2004; Kurtz & Sherker, 2003; Story, 2003). This means that observing the target acting in hundreds or even thousands of similar situations improves judgement accuracy only marginally (Kenny, 2004).

However, it is understandable that the accuracy of personality judgments would increase with the intimacy of the relationship. We can say, as it turns out, much less about friends and acquaintances compared to the information what we have about our family members or partners. Observing targets in situations that are not available to acquaintances allows judges to access new information about their personality (Ambady & Rosenthal, 1992; Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004; Kenny, 2004; Levesque & Kenny, 1993; Swann & Gill, 1997). For instance, those who have intimate relationships are more knowledgeable about thought and feelings of their targets and not only their observable behaviors (Letzring & Human,

2014). It was noticed, for example, that family members achieve stronger agreement on personality traits than those who are just friends or acquaintances (Connolly et al., 2007). Typically, spouses and partners spend a lot of time together in very different situations, they become familiar with each other's thoughts and feelings, and get used to reactions that are not necessarily available to everyone (South, Oltmanns, Johnson, & Turkheimer, 2011). Correlations between spouses' ratings of personality are often higher than correlations between family members, friends, or acquaintances (Beer & Watson, 2008; De Vries, Lee, & Ashton, 2008; McCrae et al., 2004). Social Investment Theory (Roberts, Wood, & Smith, 2005) posits that normative personality traits develop through investment in social institutions, such as age-graded social roles. It is obvious that the roles of friends, family members, and romantic partners in personality judgment differ by the type of situation witnessed and the amount of information accumulated. Even if information about personality dispositions is mostly in the public domain, different facets of personality are expressed differently across social roles. If there is a chance of finding the elusive moderators of self-other agreement, then level of relational closeness or intimacy is perhaps the most promising place to look for it.

One reason why it is difficult to find individual differences in self-other agreement is that the task of judging personality is a relatively simple one, in most cases at least. Perhaps not perfectly, but most healthy people can judge whether someone is talkative or reticent, friendly or reserved, vulnerable or impervious, dependable or unreliable. Not all mathematically or computationally complex attributes are necessarily difficult to judge. For example, analytically it is a tedious or seemingly entirely impossible mission to establish the area or volume of irregular shapes or bodies. Nevertheless, Archimedes solved such a problem posed to him by Hiero of Syracuse by immersing an irregular shaped body into a vessel filled with fluid. Similarly,

German mechanic Jakob Amsler invented a disarmingly simple device—the planimeter—to measure the area of irregular shapes like pieces of land on a map (Runeson, 1977). Considering the topic of this paper, even if personality structure turns out to be very complex, individuals may nevertheless possess smart mechanisms or at least simple heuristics which help them to evaluate not only their own but also other people's personality traits. Judgment of personality may not require sophisticated skills, which involve diligence or prolonged practice to achieve perfection. For instance, it is likely that only a tiny fraction of people has enough talent to learn to play the violin well or to juggle five balls simultaneously. Personality judgments may not be like these sophisticated skills. They may more closely resemble the dexterity needed to ride a bike, a task that most people can indeed accomplish (notwithstanding, of course, the falls and resultant bruises in the early stages of learning). There is no doubt, however, that some judgments concerning other people are difficult to make. For example, research suggests that most people, irrespective whether they are laypeople or professionals in the detection of lying, cannot tell from an individual's demeanor if she or he is telling the truth or not (Ekman, O'Sullivan, & Frank, 1999). In detecting lying, judges range no more widely than would be expected by chance, and the best judges are no more accurate than a stochastic mechanism (Bond & DePaulo, 2006, 2008). Nevertheless, it seems to be relatively easy to judge the personality traits; at least those that are usually included in popular omnibus personality questionnaires. Indeed, it is not very complicated to tell the difference between people who are, relative to an average, outspoken or reserved, emotional or unemotional, hardworking or lackadaisical, open or closed, agreeable or antagonistic, or fair or unfair, which are the typical dispositions measured by personality questionnaires.

A series of demonstrations in which ordinary participants could quite accurately describe the personality of those whom they had never met before provided evidence that personality judgment is easier than initially thought. It takes only 5-10 minutes and few dozen observable acts (e.g., making a comment, citing a proverb, presenting a request, etc.) to be able to make reasonably accurate personality judgments (Albright, Kenny, & Malloy, 1988; Borkenau & Liebler, 1992; Kenny, Horner, Kashy, & Chu, 1992; Levesque & Kenny, 1993; Realo et al., 2003; Shevlin, Walker, Davies, Banyard, & Lewis, 2003; Zebrowitz & Collins, 1997). What is more, ordinary laypeople can deduce personality traits with considerable accuracy based on the strength of a handshake (Bernieri & Petty, 2011; Chaplin, Phillips, Brown, Clanton, & Stein, 2000), how people smell (Sorokowska, Sorokowski, & Szmajke, 2012), how they say 'hello' (McAleer, Todorov, & Belin, 2014), and what their offices and bedrooms look like (Gosling, 2008; Gosling, Ko, Mannarelli, & Morris, 2002).

Another line of evidence suggesting that personality judgments may be less demanding than initially thought comes from meta-analyses on observer accuracy. The main conclusion of these analyses is that, for most personality traits, targets and informants tend to achieve self-other agreement in the range of .40 – .50 (Connelly & Ones, 2010; Connolly et al., 2007; Kenny & West, 2010; Watson, Hubbard, & Wiese, 2000). For example, a meta-analysis which integrated findings based on 44,178 target individuals rated across 263 independent samples estimated mean corrected self-other correlations at .40 (Neuroticism), .51 (Extraversion), .39 (Openness to Experience), .40 (Agreeableness), and .44 (Conscientiousness) (Connelly & Ones, 2010). These values could have even been higher if incidental acquaintances and strangers had been excluded. Median cross-observer trait agreement in cross-cultural studies was found to be .45, .48, .46, .36, and .44 for Neuroticism, Extraversion, Openness, Agreeableness, and Consciousness,

respectively (McCrae et al., 2004). Almost identical levels of self-other agreement were found in North American samples and those collected elsewhere (McCrae et al., 2004).

These reported values are the mean trait correlations. Trait correlations cannot tell, allegedly at least, the percentage of self-informant pairs who reach or fail to reach agreement in their judgments. However, mean trait (across pairs of judges) and profile correlations (across traits) are basically the same measures of agreement (Allik, Borkenau, Hrebícková, Kuppens, & Realo, 2015). In this study, it was possible to prove that if normativeness is separated from distinctiveness of personality scores and individual profiles are ipsatized, then mean profile agreement becomes identical to mean trait agreement. Analogously, we can always decompose agreement between two personality profiles into the individual contributions of traits from which these profiles are assembled. This explains why mean profile correlations are typically also in the range of .40 to .50, or even slightly higher (Allik, Realo, Mõttus, & Kuppens, 2010; Borkenau & Zaltauskas, 2009; Dobewall, Realo, Allik, Esko, & Metspalu, 2013; Kenny & Winquist, 2001; McCrae, 2008; Pelham, 1993). Although these values are clearly above the level of statistical significance, they are less impressive when we consider that correlations between two randomly paired profiles can be in the same range.

The Present Study

The main aim of this study is to examine how easy and widespread accurate personality judgments are. We investigate self-other agreement in two large samples (total N = 11,096) of participants who speak two different languages, Dutch and Estonian, and completed two different personality questionnaires, the *NEO PI-3* and *HEXACO-PI-R*. Millions of people worldwide take personality tests each year. Judging one's own or somebody else's personality has become a routine part of everyday life, especially for those involved education or

employment counseling. Selling personality tests and interpreting the results are a multi-million business for the testing industry. Besides commercial use, academic psychologists have, in recent years, been able to collect personality data from hundreds of thousands of participants in individual countries (Obschonka, Schmitt-Rodermund, Silbereisen, Gosling, & Potter, 2013) or, indeed, from all over the world (Gebauer et al., 2015; Lippa, 2010) by using tools provided by the Internet. Thus, personality judgment has become a part of everyday life, and is not just a curiosity of psychological laboratory experiments, which are, in any case, known to be prone to bias and error (Robbins & Krueger, 2005). Our purpose is to demonstrate that if we choose two people who know each other well then they, as a rule, can reach remarkably accurate personality judgments of the target person. In doing so, we also demonstrate that the correlation observed between self- and other-rated personality profiles is clearly above that observed between two randomly paired profiles, in other words, the level of chance. We argue that an outstanding level of agreement between self and other personality judgment is achieved mainly because personality judgment is a mundane task, which can be accomplished relatively easily based on available information without the need for sophisticated cognitive skills. Because personality judgments can be shown to be effortless, no moderators can seriously enhance or obstruct these judgments.

Methods

Estonian Sample

Participants in the Estonian sample came from the Estonian Biobank cohort, whose data were collected by the Estonian Genome Centre (EGC) of the University of Tartu (Leitsalu et al., 2014). Participants were recruited on a voluntary basis from among the Estonian resident adult

population (aged over 18 years). The current number of participants—close to 52,000—represents a large proportion, approximately 5%, of the Estonian adult population. The age structure of the sample is well matched to the age structure of the entire population. A portion of the participants who donated blood samples for DNA were also asked to answer personality questionnaires. Some of these data have been analyzed in previous publications (Allik, Borkenau, Hrebickova, Kuppens, & Realo, 2015; Allik, Realo, Mõttus, Esko, et al., 2010; Rene Mõttus et al., 2015). Personality data were collected from a smaller subsample of the gene donors. After removing incomplete data, there were 3,345 participants (1,984 females and 1,361 males) in the Estonian sample with a mean age of 46.4 years (SD_{age} = 17.0, ranging from 18 to 91 years) who completed the self-report of the Estonian version of the NEO Personality Inventory-3 (NEO PI-3; McCrae, Costa, & Martin, 2005). The Estonian version of the NEO PI-3 is a slightly modified version of the NEO PI-R (P. T. Costa, Jr. & McCrae, 1992; Kallasmaa, Allik, Realo, & McCrae, 2000). Like the original NEO PI-R, the NEO PI-3 has 240 items that measure 30 personality facets, which are grouped into the five FFM domains – Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C) – such that each domain score is a composite of six facet scores. The items are answered on a five-point scale (0 = *false/strongly disagree* to 4 = *true/strongly agree*).

All 3,345 participants nominated somebody who knew them well and these informants were asked to rate the personality of the participant using the other-report version of the Estonian NEO PI-3. Of the informants, 2,331 were female (71.1%) and 948 were male (66 did not report their gender). The mean age of informants was 41.8 (*SD* = 15.9) years. The informant questionnaire also contained several questions about the relationship to the target, asking, for

example, how long they had known the target. The mean length of acquaintance was 23.2 years (*SD* = 15.1).

Dutch Sample

The Dutch sample consisted of 2,198 first year personality psychology students (81.7% female; M_{age}=20.2; SD_{age}=2.8), who, as part of their course work, filled out personality inventories and recruited a well-acquainted other to obtain informant-reports on their personality. The sample of well-acquainted informants consisted of 41.2% friends, 35.9% family members, and 22.9% intimate partners, who, on average, had known the target person for 10.6 years (*SD*=8.0). Parts of the dataset have been used in other studies (e.g., De Vries, Ashton, & Lee, 2009; De Vries, Wawoe, & Holtrop, in press, Studies 1 and 4). Both self- and other-ratings were obtained using the Dutch 200-item HEXACO Personality Inventory-Revised (HEXACO-PI-R), which measures six personality domains – Honesty-Humility (H), Emotionality (E), Extraversion (X), Agreeableness (A), Conscientiousness (C), and Openness (O) –each of which is comprised of four subscales (de Vries, Ashton, & Lee, 2009; De Vries et al., 2008). In total, HEXACO-PI-R thus has 24 facet scales, each of which is, in turn, measured by 8 items. In addition, there is a 25th facet scale – Altruism – which is interstitial in the personality space, spanning Honesty-Humility, Emotionality, and Agreeableness. In order to make HEXACO-PI-R profiles maximally comparable to those of NEO PI-3, we used the first 24 facet scales in the analysis. For convenience, we also changed the usual order of the HEXACO-PI-R facet scales, arranging them as NEOA(H)C, which is a more familiar order for NEO PI-3 users.

Results

Trait-by-trait correlation. Firstly, we computed trait-by-trait correlations for all facet scales. Figure 1 demonstrates the correlations for the facets of NEO PI-3 (A) and HEXACO-PI-R (B).

The average self-other correlations across all facets were .46 and .49 for the Estonian and Dutch samples, respectively ($p < .0001$). These two values are in the typical range for self-other agreement reported in the literature (Connelly & Ones, 2010; De Vries et al., 2008; McCrae et al., 2004).

In both samples and instruments, a substantial fluctuation in the level of agreement can be observed. For example, for HEXACO-PI-R facets, the lowest self-other agreement was found for H1: Sincerity ($r = .31$) and the highest for C1: Organization ($r = .66$). Even from this example, it is difficult to see why a tendency to be genuine in interpersonal relations is so much less visible or difficult to judge than a tendency to seek order, particularly in one's physical surroundings. Self-other agreement was higher in the subscales in which individual differences were larger (cf.; Allik, Realo, Mõttus, Esko, et al., 2010). Indeed, self-other trait agreement correlated significantly with the combined self and other standard deviation of mean scores: $r = .71$ and .51, respectively, for NEO PI-3 and HEXACO-PI-R.

Thus, differences in self-other trait correlation may be caused by a restricted variance. Assuming that the unrestricted variance for all subscales is equal to the maximally observed variance, it would be possible to calculate the correction for the observed self-observer correlation according to the formula known as the Thorndike Case 2 (Thorndike, 1949). The red broken line in Figure 1 shows self-other agreement corrected for restricted variance separately for NEO PI-3 and HEXACO-PI-R.

Insert Figure 1 about here

After the correction, the standard deviation of self-other correlations was slightly reduced but not enough for a claim that all differences in self-other correlation were eliminated.

However, several changes in correlation were not in the direction that would be expected from the trait observability. For example, before the correction one of the lowest agreements was for the NEO PI-3's O6: Values. After correcting for restricted variance it jumped from .37 to .55 which was above the average. Thus, a relatively low self-other correlation on openness to values was not because external judges cannot observe someone's readiness to reexamine his or her social, political, and religious values but mainly because there was not enough variation between respondents' answers.

Correlation between profiles. Apart from a trait-by-trait approach, it is also possible to compute the correlation between two profiles of K traits (K = 30 for the NEO PI-3 and K = 24 for the HEXACO-PI-R). Although it is believed that trait-by-trait correlations have a substantively different meaning to correlations between two profiles, they are mathematically strongly related. Moreover, when done on both normalized and ipsatized scores, they are identical (Allik, Borkenau, Hrebícková, et al., 2015). Figure 2 demonstrates the histograms of the self-other profile correlations for the Estonian and Dutch samples.

Insert Figure 2 about here

The mean correlations were .613 and .595 for NEO PI-3 and HEXACO-PI-R, respectively, which is, expectedly, higher than the mean trait-by-trait correlations (Allik, Borkenau, Hrebícková, et al., 2015). Although a correlation of about .60 may seem impressive, it is less so if the correlation between two randomly paired personality profiles is not zero, but falls, say, somewhere close to .40, leaving a very narrow margin between genuine accurate perception and that which occurs by chance.

In order to decide whether the mean profile correlations .613 and .595 are sufficiently above the level of chance, we need to establish a distribution of correlation coefficients between randomly paired personality profiles. One possible way of constructing this distribution is to pair each self-rated personality profile with every *N*-1 or "wrong" other-rated profile in the sample. Thus, in each sample, along with *N* true self-other pairs, we can form *N*·(*N*-1) "wrong" pairs of profiles between two arbitrarily connected self- and other-rated profiles. For each target, we can compute the mean and standard deviation correlations between *N*-1 arbitrarily linked profiles. If we subtract the mean correlation of the arbitrary pairings from the true self-other correlation, and divide the obtained difference by the standard deviation of the arbitrary pairings, then we have the difference between true and arbitrary pairings in the units of the standard normal distribution.

The mean correlation between two arbitrarily paired profiles was .281 and .198 for NEO PI-3 and HEXACO-PI-R samples, respectively (see vertical lines in Figure 1). Thus, the mean differences between profiles were reduced to .332 and .397, respectively. If we divide these values by the respective standard deviations of the correlations of the arbitrary pairings, then we convert these values into units of the standard normal distribution, or *z*-values. The z-values were 1.395 and 1.724 for NEO PI-3 and HEXACO-PI-R samples, respectively. This means that in 91.8% and 95.8% of all cases, for Estonian and Dutch samples, respectively, the correlation between the profiles of true pairs is higher than the correlation between the personality profiles of two randomly paired individuals.

The fact that the true profile correlation can be confused with the correlation of two randomly paired profiles in only 4-8% of cases is remarkable. It is remarkable because the raw profile scores contain distinctive information—how much the individual is above or below average scores on each trait—in addition to normative information, or how much the individual

is similar to an average person, in general (Cronbach & Gleser, 1953; Gage & Cronbach, 1955). In order to separate distinctiveness scores from normative scores (the degree to which a profile reflects an average profile), all data were normalized by transforming them into T-scores (with a mean equal to 50 and a standard deviation equal to 10). The Estonian sample was standardized in four separate groups of participants (males and females, each divided into two groups: one younger and one older than 30 years). Since participants in the Dutch sample were almost exclusively younger than 30 years, their data were standardized by target sex only. After standardization, the profile correlations dropped substantially, as was expected. In the Estonian sample, the mean correlation between profiles dropped from .613 to .444, and in the Dutch student sample from .595 to .473. Because the mean correlation between two randomly paired distinctive or t-scored profiles is close to zero, the distinctiveness of the true correlation from random pairing becomes even greater. The mean difference between true and random correlations in *z*-scores was 1.56 and 1.782 for NEO PI-3 and HEXACO-PI-R samples, respectively. Translating these scores into probabilities, we can state that true profile pairs can be confused with an arbitrary pairing in only 5.9% and 3.7% of all cases for NEO PI-3 and HEXACO-PI-R samples, respectively. This means that self-other agreement is above the level of chance in 94%-96% of cases.

Moderator variables. Among potential moderator variables, age of targets is perhaps one of the most interesting. For example, it is logical to expect that a young person is more difficult to judge than a person of advanced age, whose personality is no longer likely to change much. Figure 3 (A: Estonian sample and B: Dutch sample) demonstrates the extent to which distinctive self-other profile agreement depends on target age. Although personality traits change with age (especially at a younger age) (Allik, Laidra, Realo, & Pullmann, 2004), the agreement between

self- and other-ratings remained at the same level. These near-zero correlations were not significant. Based on these negative findings, we can conclude that age of targets plays no prominent role in the agreement between judges (cf.; McCrae & Costa, 1982). A mature person's personality is no more judgeable than that of a young individual, whose personality is still fluid and prone to changes.

Insert Figure 3 about here

Next, we analyzed the impact of gender on self-other agreement. According to popular belief, women are more empathically accurate than men (cf.; Klein & Hedges, 2001; Laurent & Hodges, 2009). If true, this would mean that women are also better judges of another person's personality. We analyzed all target-judge gender combinations: female-female, female-male, male-female, and male-male, where the first position denotes target, and the second, judge. Figure 4 demonstrates the mean profile correlations for these four target-judge combinations.

Insert Figure 4 about here

Figure 4 demonstrates that women are not superior to men in judging others' personality. In general, the gender of targets and judges is irrelevant. In the Dutch sample, an ANOVA gave a nonsignificant result: $F(3,2194)=1.83$, $p = .140$, $\eta^2 = .002$. Although, in the Estonian sample, the gender combination of target-judge pairs was statistically significant – $F(3,3275) = 4.23$, $p = .005$, $\eta^2 = .004$ – the effect size η^2 (eta-square) explained only about 0.4% of the total variance. In sum, there appears to be a small advantage when judging personality in opposite gender dyads, but it is doubtful whether such a small effect has any theoretical or practical significance.

Whether length of acquaintance affects self-other agreement has been disputed for quite some time. For example, according to some estimates, the level of self-other agreement increases about $\Delta r = .05$ for every 5 years of acquaintance (Biesanz et al., 2007). Informants in our two samples reported that they had known the target for an average of 21.3 and 10.6 years in the Estonian and Dutch samples, respectively. There was no correlation between length of acquaintance and agreement in the NEO PI-3 sample: $r(3277) = .02$, $p = .256$. However, in the Dutch sample, there was a minor positive correlation between length of acquaintance and self-other agreement: $r(2198) = .048$, $p = .023$. Nevertheless, from both theoretical and practical viewpoints, an increase in agreement $\Delta r = .01$ for every 10 years of acquaintance does not seem very meaningful. If there is an improvement with length of acquaintance, then it happens in the very early stages of acquaintance, after which a plateau is reached (cf., Biesanz et al., 2007; Gnambs, 2013; L. Schneider, Schimmack, Petrican, & Walker, 2010).

Relationship. How does the type of the relationship between the judge and the target influence self-other agreement in personality? In both samples, the relationship between judge and target was classified into three categories: (1) friends and acquaintances, (2) family members (parents, children, and siblings), and (3) partners (including spouses). These three categories were arranged in order according to the mean scores in response to the statement "I know this person well." In both samples, informants stated that they knew members of their family better than friends or acquaintances. Similarly, informants from both the Dutch and Estonian samples thought that, on average, they knew their partners or spouses better than they knew members of their family. An ANOVA confirmed this association at an extremely high level of significance: $F(2,3251) = 1472.95$, $p < .00001$, $\eta^2 = .475$ and $F(2,2195) = 199.65$, $p < .00001$, $\eta^2 = .154$ respectively, for Estonian and Dutch samples. Figure 5 demonstrates the mean profile correlation

for the three different categories of relationship between judge and target – friends, family, and partners. As expected, the observed differences in the similarity of profiles were highly significant. An ANOVA revealed that spouses or partners reached a much better agreement than members of family who, in turn, achieved a greater consensus than friends or acquaintances: $F(2,3280) = 9.28$, $p = .00001$, $\eta^2 = .006$. The same pattern was observed in the Dutch sample: $F(2,2195) = 18.59$, $p < .00001$, $\eta^2 = .017$. The observed pattern is clear: the more closely or intimately you know the target, the stronger the agreement between self and informant personality judgments.

Insert Figure 5 about here

Discussion and Conclusions

Although no description of a person is possible without trait attributes, it is also believed that traits themselves yield little beyond what has been called the "psychology of the stranger" (McAdams, 1995). Just as all cats are gray in the dark, people lack individuality when described by a universal set of personality dispositions. According to this point of view, trait attributes used when judging personality are too impersonal to capture the uniqueness of each individual. In this study, for the first time, we constructed a statistical distribution with which we were able to calculate how often agreement between self- and other-rated personality profiles is confused with agreement occurring from a random pairing of two unrelated profiles. The results from two large samples, one Dutch and one Estonian, and two conceptually different personality instruments, NEO PI-3 and HEXACO-PI-R, demonstrated that in no more than 4-6% of cases could true agreement be confused with random agreement between two unrelated profiles. That is, in the

vast majority of cases, a true agreement between self-rated and other-rated profiles could be achieved. Two clear implications deserve mention at this point. First, individual profiles across universal sets of personality dispositions are unique, such that one individual can be distinguished from others. Second, a judge who knows her or his target well can describe this unique disposition profile remarkably accurately and far beyond the level of randomly occurring agreement.

In spite of the unusual power we had in this study (the total number of participants was over 11,000), the search for a variable, which could seriously moderate self-other agreement was largely unsuccessful. Just as in many previous studies, we were not able to identify a distinctive group of people who are "good targets" for personality judgments of knowledgeable others. For example, it did not matter whether target persons' personality was still in development or already "set like plaster" by 30 years of age as William James argued in 1890 (James, 1890, p. 105) – self-other agreement was effectively at the same level (for a debate of the plaster hypothesis see P. T. Costa & McCrae, 2006; Roberts, Walton, & Viechtbauer, 2006). We also failed to identify who "good judges" are. As an example, in spite of popular lore that women are generally better judges of personality, we found men equally capable of judging another person's personality. We also did not find a coherent grouping of traits on which agreement is slightly better than on others. Some researchers have proposed distinguishing surface traits, which are more amenable for external inspection, from other traits which are more deeply buried within the person (Asendorpf & Van Aken, 2003; Epstein, 2010; Kandler, Zimmermann, & McAdams, 2014; Swider, Zimmerman, Charlier, & Pierotti, 2015). However, as variation in agreement is relatively modest, this distinction would not seem to hold. For example, it was easier for participants to agree on Fearfulness and Depression, which are facets of Neuroticism, than on

Impulsiveness and Sentimentality, which are also manifestations of neurotic tendencies. It is also possible that achieving agreement on traits related to moral character is more complicated than on more observable traits but our data provided no convincing evidences in favor of this proposal (cf., Cohen, Panter, Turan, Morse, & Kim, 2013).

Two judges seem to achieve at least satisfactory agreement on virtually all personality traits included in the major questionnaires, but this agreement may be even stronger on some traits. It is not immediately apparent how visibility or social desirability could lead to this pattern of agreement (Allik, Realo, Mõttus, Borkenau, et al., 2010). Personality traits belonging to the one dimension may have very different levels of visibility, indicating that this is not a consistent attribute of personality traits. Analogously, the mean values of the facet scales belonging to the same dimension may have very different standard deviations. Nevertheless, the range of interindividual variance in facet scores explains the largest portion of variance in the level of self-other agreement (see also Allik, Realo, Mõttus, Esko, et al., 2010). Thus, differences in the level of agreement are mainly "technical," depending on a selection of appropriate items that allow for sufficient individual variation in answers. There is no good explanation why, for example, the standard deviation of O2: Aesthetics, a facet scale of the NEO PI-3, is about two times higher than that of O6: Ideas. It is hard to believe that variability in a deep appreciation for art and beauty—the main content of the O2: Aesthetics—is inherently larger than variability in the readiness to re-examine social and political values—the main content of the openness to new ideas. This is, thus, an unresolved issue, which needs to be studied further.

It could be argued that our failure to find moderators of self-other agreement is an exaggeration or an overstatement. Indeed, nobody can say that there is no evidence in favor of "good target", "good judge", or "good information." All what we are saying is that despite some

very interesting and important findings, there still is not a clear answer to the question of who these good targets and judges are (Human & Biesanz, 2013, p. 248). If we accept a conclusion that moderator effects in personality judgments are small and can only be detected in very large samples (Chaplin, 1991) then we may need a more general explanation for a failure to find moderators effect. By any means this study is not conclusive but a cautious attempt to provide a tentative answer to a question why moderators of self-other agreement are difficult to establish.

One possible explanation for why it was so difficult to find reliable individual differences or other moderators of the accurate personality judgments is that judging another person's personality is a rather trivial task, which does not need special practice or sophisticated social skills. After analyzing evidences, it seems that the complexity of personality judgments has been overestimated. Instead of the complexity thesis, it is more plausible to talk about the fundamental simplicity of personality judgments. Usually, people know a lot about the personality—the way in which they think, feel, and act in a variety of everyday situations—of those with whom they interact on a daily basis. It is correct, of course, that the person being accurately judged must exhibit his or her traits, which must be available to the judge, who must register this information and make valid conclusions on its basis (Funder, 2012). However, this study demonstrates, along with many similar findings, that if we more or less randomly take two people from the general population who know each other sufficiently well, in most cases, a level of consensus in judging personality traits, which is considerably above chance, is achieved. This means that, in most cases, trait relevant information was available, which was registered, processed, and adequately used to make valid trait inferences.

What our data seem to suggest is that our daily life contains an abundance of information about personality traits. Nobody can live a normal life without constantly emanating relevant

information about his or her personality dispositions. We can refer here to a vast amount of literature on so-called zero acquaintances and thin slices of experience (Albright et al., 1988; Ambady & Rosenthal, 1992; Beer & Wayson, 2008; Borkenau & Liebler, 1992; Borkenau et al., 2004; Hirschmüller, Egloff, Schmukle, Nestler, & Back, 2015; Leikas, Verkasalo, & Lönnqvist, 2013; Tackett, Herzhoff, Kushner, & Rule, 2016). All these data, together with the present study, seem to indicate that personality judgments are not very sophisticated, and that they do not require huge amounts of practice to be carried out successfully. We are not aware of any study in which at least moderate self-other agreement was not achieved on the majority of personality traits studied (Connelly & Ones, 2010; Connolly et al., 2007; Kenny & West, 2010; McCrae et al., 2004; Watson et al., 2000). This means that lay perceivers have achieved accuracy in personality judgement more than once in the lifetime. If somebody laughs frequently and talks a lot, then it does not require above average intelligence to conclude that this person might be an extravert. If a person is consistently critical and antagonistic, it is very likely that he or she would receive rather low scores on agreeableness. If somebody we know is never late for appointments and he or she keeps things in perfect order, then we are likely to consider him or her to be conscientious. There is no doubt that these simple conclusions require data as well as some basic logic. Judgment of personality requires simple skills and competencies that the majority of the population possesses. Without these skill and competencies, it would be very complicated to function normally in social environments (Haselton & Funder, 2006).

Although judging another person's personality is not a complicated task and is one in which most people appear to have around the same level of skill, a close relationship between judge and target, nevertheless, seems to make agreement stronger: a relationship that is more intimate seems to make personality judgments more accurate. This pattern replicates findings in the

previous literature. Relatives and family members usually achieve stronger self-other agreement than those who do not know their targets as well (Connelly & Ones, 2010; Connolly et al., 2007). It is also true that spouses typically achieve higher levels of self-other correlations than siblings (Beer & Watson, 2008; McCrae et al., 2004). A general rule is that agreement between self- and informant-ratings increases with degree of acquaintanceship – that is, a close or intimate relationship increases level of agreement (Connelly & Ones, 2010; De Vries et al., 2008; Kenny, 2004; Kenny, Albright, Malloy, & Kashy, 1994; Letzring, Wells, & Funder, 2006; Paunonen, 1989; Starzyk, Holden, Fabrigar, & Macdonald, 2006). However, there is evidence that consensus between judges does not always increase with higher levels of acquaintanceship (Kenny, 2004). The effect of acquaintanceship is in accordance with the RAM, which predicts that information quantity and quality are positively related to objective knowledge about the targets, which, in turn, leads to increased levels of judgment accuracy (Beer & Brooks, 2011; Funder, 1995, 2012). As our results indicate, the quantity of information has nothing to do with the length of acquaintanceship. Some sources of information seem to be more informative than others, in spite of the fact that a huge amount of personality information is freely available. However, it does not really matter how long a person has not had access to a relevant piece of information.

Most likely, several factors contribute independently to the observed friend-family-partner agreement gradation. First, this ranking may correspond to degrees of target-judge similarity. Judges are likely to use projection, assuming that their target has a similar personality profile – assumed similarity – which can be estimated by the extent to which people judge other individuals as consistent with how they judge themselves (de Vries, 2010; Kenny & West, 2010). Thus, judges have better a chance of agreeing with targets' self-judgment if their personality

profiles are similar to some extent. For example, family members have personality profiles that are more similar to each other, compared with two genetically unrelated individuals, because of shared genes (Jang, Livesley, & Vernon, 1996; Kandler, Riemann, Spinath, & Angleitner, 2010). Married or dating couples may be more similar to each other due to assortative mating, implying that people seek mates with personality traits similar to themselves (Luo & Klohnen, 2005; McCrae et al., 2008).

If there is a deficit in relevant information, then informants have a tendency to fill the gap with assumed information. Unfortunately, our study did not ask informants to rate their own personality as well. For this reason, we cannot estimate directly assumed similarity, or the extent to which targets were described in a similar way to which judges viewed themselves. Although it is logical to assume that a shortage of information can be compensated for with assumed information (Beer & Watson, 2010; Beer, Watson, & McDade-Montez, 2013; Human & Biesanz, 2011; Kenny & Kashy, 1994; Lee et al., 2009; Paunonen & Kam, 2014; Rogers & Biesanz, 2015), a previous study has shown that assumed knowledge is almost irrelevant in self-other agreement (de Vries, 2010). In the present study, we also demonstrated that the ratio between normativeness and distinctiveness in self-other agreement remained approximately the same for a group of friends, family members, and partners. This means that we still have a very limited understanding what married or unmarried couples have that family members do not; and what family members have that friends and acquaintances do not. One possible candidate is an access to some privileged information. For example, it was found that self-other agreement was higher among judges exposed to information about thoughts and feelings of the target than for judges exposed to information about hobbies and activities (Andersen, 1984). These results were confirmed by other studies demonstrating that dyads assigned to discuss thoughts and feelings

achieved higher distinctive accuracy than participants, who were engaged in behaviors or discussing these behaviors (Letzring & Human, 2014). Thus, besides observing behaviors it is beneficial to know about thoughts, feelings, and motives that are behind the observed conducts. If this line of reasoning is true then married or unmarried couples are expected to know more about thought and feelings of their partners than family members usually do; family members, in turn, are expected to know more about one another's views and motives than friends and acquaintances usually do.

Studies have shown that observers' ratings of personality may predict performance behaviors better than do self-ratings (Connelly & Hülsheger, 2012; Kolar, Funder, & Colvin, 1996; Vazire, 2010). For example, it was proposed that the predictive advantages of other-ratings may stem from use of observers who have a frame of reference more closely aligned with the criterion ("narrower scope") or observers having greater accuracy than targets themselves ("clearer lens") (Connelly & Hülsheger, 2012). Unfortunately, the used study format – there was no performance to predict – did not allow to decide which of the two ratings, self- or other-ratings was more accurate. Nevertheless, it is possible to speculate about "narrower scope" and "clearer lens" hypotheses as potential explanations for the friend-family-partner gradient. Because the both hypotheses represent variants of the "good target" or "good judge" type of explanations, we take a liberty to be skeptical towards them. In the meantime, we can notice that the accuracy gradient corresponds to a general pattern of social relations (Fiske, 1992). It seems that the social relationships are organized on the basis of interpersonal distance from very close relations between married or unmarried couples and members of nuclear family to remote contact with strangers and impersonal orders, requests, and memos from organizations and institutions (Realo, Allik, & Vadi, 1997). If we suppose that, the role of personality dispositions is more prominent

in close or intimate relations and progressively less important in more distant relationships, it could be a potential constraint we are looking for.

Finally, it was encouraging to see how similar the data obtained from two different samples using two different personality instruments were. The participants in each sample spoke languages belonging to two different language groups: Indo-European and Finno-Ugric. The mean age of participants had a disparity of more than 21 years. The two instruments, HEXACO-PI-R and NEO PI-3, conceptualize the space of personality traits differently, by assuming that there are five independent personality dimensions on the one hand, and six on the other. Nevertheless, the basic findings were strikingly similar, and it can be almost certainly excluded that this was due to some kind of statistical fluke or method bias. In both samples, the percentage of target-judge pairs who achieved satisfactory agreement between personality profiles was very similar. In both samples, demographic information – age and sex – had a negligible moderating effect on self-other agreement. The same was true for duration of acquaintance: neither Dutch nor Estonian participants benefitted from knowing their targets for a longer period. However, intimate knowledge helped married and unmarried couples in both samples be more accurate in their personality judgments. Family members, in turn, had knowledge that made them more accurate than two people who were just acquaintances or friends, again similarly in both the Dutch and Estonian samples. The type of relationship aside, most other moderating variables (including individual differences) had a negligible effect on self-other agreement.

To conclude, because judging another person's personality is a rather simple task, it seems increasingly unlikely that researchers will be able to make a reliable distinction between "good or bad targets," "good or bad traits," and "good or bad judges."

Authors Note

Jüri Allik, Department of Psychology, University of Tartu, Estonia and the Estonian Academy of Sciences; Reinout E. de Vries, Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam and Department of Psychology, University of Twente, the Netherlands; Anu Realo, Department of Psychology, University of Warwick, UK and Department of Psychology, University of Tartu, Estonia.

Preparation of this manuscript was supported by the University of Tartu (SP1GVARENG) and by institutional research funding (IUT2-13) from the Estonian Ministry of Education and Science to Jüri Allik. Anu Realo was a Visiting Professor at the Health, Medical, and Neuropsychology Department at the Faculty of Social and Behavioral Sciences at Leiden University during the writing of this article. We are grateful to the Estonian Genome Centre of the University of Tartu and its director, Andres Metspalu, for help in collecting the data and the kind permission to use the data in the current study. The assistance of Liisi Kööts-Ausmees in organizing the data file prior to analysis is also gratefully acknowledged.

Correspondence concerning this article should be addressed to Jüri Allik, Department of Psychology, University of Tartu, Näituse 2, Tartu 50409, Estonia. Electronic mail may be sent to juri.allik@ut.ee.

References

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55*(3), 387-395. doi:10.1037/0022-3514.55.3.387

Allik, J., Borkenau, P., Hrebickova, M., Kuppens, P., & Realo, A. (2015). How are personality trait and profile agreement related? *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.00785

Allik, J., Borkenau, P., Hrebícková, M., Kuppens, P., & Realo, A. (2015). How are Personality Trait and Profile Agreement Related? *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.00785

Allik, J., Laidra, K., Realo, A., & Pullmann, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. *European Journal of Personality, 18*(6), 445-462. doi:10.1002/per.524

Allik, J., Realo, A., Mõttus, R., Borkenau, P., Kuppens, P., & Hřebíčková, M. (2010). How People See Others is Different from How People See Themselves: A Replicable Pattern across Cultures. *Journal of Personality and Social Psychology, 94*, 168-182. doi:DOI: 10.1037/a0020963

Allik, J., Realo, A., Mõttus, R., Esko, T., Pullat, J., & Metspalu, A. (2010). Variance determines self-observer agreement on the Big Five personality traits. *Journal of Research in Personality, 44*(4), 421-426. doi:10.1016/j.jrp.2010.04.005

Allik, J., Realo, A., Mõttus, R., & Kuppens, P. (2010). Generalizability of self-other agreement from one personality trait to another. *Personality and Individual Differences, 48*(2), 128-132. doi:10.1016/j.paid.2009.09.008

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt, Rinehart & Winston.

Allport, G. W. (1966). Traits revisited. *American Psychologist,, 21*(1), 1-10. doi:10.1037/h0023295

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256-274. doi:10.1037/0033-2909.111.2.256

Andersen, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology, 46*(2), 294-307. doi:10.1037/0022-3514.46.2.294

Asendorpf, J. B., & Van Aken, M. A. G. (2003). Personality-Relationship Transaction in Adolescence: Core Versus Surface Personality Characteristics. *Journal of Personality, 71*(4), 629-666. doi:10.1111/1467-6494.7104005

Ashton, M. C., Lee, K., de Vries, R. E., Perugini, M., Gnisci, A., & Sergi, I. (2006). The HEXACO model of personality structure and indigenous lexical personality dimensions in Italian, Dutch, and English. *Journal of Research in Personality, 40*(6), 851-875. doi:10.1016/j.jrp.2005.06.003

Bayne, R. (1985). The good judge of personality: A review. *Bulletin of the British Psychological Society, 38*(FEB), A46-A46. Retrieved from <Go to ISI>://WOS:A1985ACT1200242

Beer, A., & Brooks, C. (2011). Information quality in personality judgment: The value of personal disclosure. *Journal of Research in Personality, 45*(2), 175-185. doi:10.1016/j.jrp.2011.01.001

Beer, A., & Watson, D. (2008). Asymmetry in judgments of personality: Others are less differentiated than the self. *Journal of Personality, 76*(3), 535-559. doi:10.1111/j.1467-6494.2008.00495.x

Beer, A., & Watson, D. (2010). The effects of information and exposure on self-other agreement. *Journal of Research in Personality, 44*(1), 38-45. doi:10.1016/j.jrp.2009.10.002

Beer, A., Watson, D., & McDade-Montez, E. A. (2013). Self-Other Agreement and Assumed Similarity in Neuroticism, Extraversion, and Trait Affect: Distinguishing the Effects of Form and Content. *Assessment, 20*(6), 723-737. doi:10.1177/1073191113500521

Beer, A., & Wayson, D. (2008). Personality judgment at zero acquaintance: Agreement, assumed similarity, and implicit simplicity. *Journal of Personality Assessment, 90*(3), 250-260. doi:10.1080/00223890701884970

Bernieri, F. J., & Petty, K. N. (2011). The influence of handshakes on first impression accuracy. *Social Influence, 6*(2), 78-87. doi:10.1080/15534510.2011.566706

Bernieri, F. J., Zuckerman, M., Koestner, R., & Rosenthal, R. (1994). Measuring person perception accuracy: Another look at self-other agreement. *Personality and Social Psychology Bulletin, 20*(4), 367-378. doi:10.1177/0146167294204004

Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self-other agreement in judgments of personality. *Journal of Personality and Social Psychology, 92*(1), 119-135. Retrieved from <Go to ISI>://000243249300010

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214-234. Retrieved from <Go to ISI>://WOS:000240159400002

Bond, C. F., Jr., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*(4), 477-492. doi:10.1037/0033-2909.134.4.477

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62*(4), 645-657. Retrieved from <Go to ISI>://A1992HM48500010

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology, 86*(4), 599-614. Retrieved from <Go to ISI>://000220317200006

Borkenau, P., & Zaltauskas, K. (2009). Effects of Self-enhancement on Agreement on Personality Profiles. *European Journal of Personality, 23*(2), 107-123. doi:10.1002/per.707

Chaplin, W. F. (1991). The next generation of moderator research in personality psychology. *Journal of Personality, 59*, 143-178.

Chaplin, W. F., Phillips, J. B., Brown, J. D., Clanton, N. R., & Stein, J. L. (2000). Handshaking, gender, personality, and first impressions. *Journal of Personality and Social Psychology, 79*(1), 110-117. Retrieved from 10909881

Christiansen, N. D., Wolcott-Burnam, S., Janovics, J. E., Burns, G. N., & Quirk, S. W. (2005). The good judge revisited: Individual differences in the accuracy of personality judgments. *Human Performance, 18*(2), 123-149. doi:10.1207/s15327043hup1802_2

Cohen, T. R., Panter, A. T., Turan, N., Morse, L., & Kim, Y. (2013). Agreement and similarity in self-other perceptions of moral character. *Journal of Research in Personality, 47*(6), 816-830. doi:10.1016/j.jrp.2013.08.009

Connelly, B. S., & Hülsheger, U. R. (2012). A Narrower Scope or a Clearer Lens for Personality? Examining Sources of Observers' Advantages Over Self-Reports for Predicting Performance. *Journal of Personality, 80*(3), 603-631. doi:10.1111/j.1467-6494.2011.00744.x

Connelly, B. S., & Ones, D. S. (2010). An Other Perspective on Personality: Meta-Analytic Integration of Observers' Accuracy and Predictive Validity. *Psychological Bulletin, 136*(6), 1092-1122. doi:10.1037/a0021212

Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment, 15*(1), 110-117. Retrieved from <Go to ISI>://000244740900012

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*: Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (2006). Age Changes in Personality and Their Origins: Comment on Roberts, Walton, and Viechtbauer (2006). *Psychological Bulletin, 132*(1), 26-28. Retrieved from ISI:000234946500002

Cronbach, L. J., & Gleser, G. (1953). Assessing similarity between profiles. *Psychological Bulletin, 50*(6), 456-473. doi:10.1037/h0057173

Davis, M. H., & Kraus, L. A. (1997). Personality and empathic accuracy. In W. Ickes (Ed.), *Personality and empathic accuracy* (pp. 144-168). New York: Guilford Press.

De Raad, B., Barelds, D. P. H., Timmerman, M. E., De Roover, K., Mlacic, B., & Church, A. T. (2014). Towards a Pan-cultural Personality Structure: Input from 11 Psycholexical Studies. *European Journal of Personality, 28*(5), 497-510. doi:10.1002/per.1953

de Vries, R. E. (2010). Lots of Target Variance: An Update of SRM Using the HEXACO Personality Inventory. *European Journal of Personality, 24*(3), 169-188. doi:10.1002/per.764

de Vries, R. E., Ashton, M. C., & Lee, K. (2009). The six most important personality dimensions and the HEXACO Personality Inventory. *Gedrag & Organisatie, 22*(3), 232-274. Retrieved from <Go to ISI>://WOS:000270277600004

De Vries, R. E., Lee, K., & Ashton, M. C. (2008). The dutch HEXACO personality inventory: Psychometric properties, self-other agreement, and relations with psychopathy among low and high acquaintanceship dyads. *Journal of Personality Assessment, 90*(2), 142-151. doi:10.1080/00223890701845195

Dobewall, H., Realo, A., Allik, J., Esko, T., & Metspalu, A. (2013). Self-Other Agreement in Happiness and Life-Satisfaction: The Role of Personality Traits. *Social Indicators Research, 114*(2), 479-492. doi:10.1007/s11205-012-0157-y

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science, 10*, 263-270.

Epstein, S. (2010). The Big Five Model: Grandiose Ideas About Surface Traits as the Foundation of a General Theory of Personality. *Psychological Inquiry, 21*(1), 34-39. doi:10.1080/10478401003648682

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*(99), 689-723.

Funder, D. C. (1995). On the accuracy of personality judgment: a realistic approach. *Psychological Review, 102*(4), 652-670. doi:10.1037//0033-295X.102.4.652

Funder, D. C. (1999). *Personality judgement: a realistic approach to person perception*. San Diego: Academic Press.

Funder, D. C. (2012). Accurate Personality Judgment. *Current Directions in Psychological Science, 21*(3), 177-182. doi:10.1177/0963721412445309

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52*(2), 409-418.

Gage, N. L., & Cronbach, L. J. (1955). Conceptual and methodological problems in interpersonal perception. *Psychological Review, 62*(6), 411-422. doi:http://dx.doi.org/10.1037/h0047205

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. doi:10.1037/pspp0000052

Gnambs, T. (2013). The Elusive General Factor of Personality: The Acquaintance Effect. *European Journal of Personality, 27*(5), 507-520. doi:10.1002/per.1933

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

Gosling, S. D. (2008). *Snoop : what your stuff says about you*. New York: Basic Books.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*(3), 379-398. Retrieved from 11902623

Haselton, M. G., & Funder, D. C. (2006). The evolution of accuracy and bias in social judgment. In M. Schaller, J. A. Simpson, & D. T. Kenrick (Eds.), *Evolution and social psychology* (pp. 15-38). New York, NY: Psychology Press.

Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D. (2015). Accurate Judgments of Neuroticism at Zero Acquaintance: A Question of Relevance. *Journal of Personality, 83*(2), 221-228. doi:10.1111/jopy.12097

Human, L. J., & Biesanz, J. C. (2011). Target Adjustment and Self-Other Agreement: Utilizing Trait Observability to Disentangle Judgeability and Self-Knowledge. *Journal of Personality and Social Psychology, 101*(1), 202-216. doi:10.1037/a0023782

Human, L. J., & Biesanz, J. C. (2013). Targeting the Good Target: An Integrative Review of the Characteristics and Consequences of Being Accurately Perceived. *Personality and Social Psychology Review, 17*(3), 248-272. doi:10.1177/1088868313495593

Ickes, W. J. (1997). *Empathic accuracy*. New York: Guilford Press.

Ickes, W. J., Buysse, A., Pham, H., Rivers, K., Erickson, J. R., Hancock, M., . . . Gesn, P. R. (2000). On the difficulty of distinguishing "good" and "poor" perceivers: A social relations analysis of empathic accuracy data. *Personal Relationships, 7*(2), 219-234. doi:10.1111/j.1475-6811.2000.tb00013.x

Ickes, W. J., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships, 7*(1), 95-110. doi:10.1111/j.1475-6811.2000.tb00006.x

James, W. (1890). *Principles of psychology*. New York: Holt.

Jang, K. L., Livesley, W. J., & Vernon, P. A. (1996). Heritability of the big five personality dimensions and their facets: A twin study. *Journal of Personality, 64*(3), 577-591. Retrieved from <Go to ISI>://A1996VD76400002

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality, 2*, 171-203.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61*(4), 521-551. Retrieved from ISI:A1993MY59000006

Kallasmaa, T., Allik, J., Realo, A., & McCrae, R. R. (2000). The Estonian version of the NEO-PI-R: An examination of universal and culture-specific aspects of the Five-Factor Model. *European Journal of Personality, 14*, 265-278.

Kandler, C., Riemann, R., Spinath, F. M., & Angleitner, A. (2010). Sources of Variance in Personality Facets: A Multiple-Rater Twin Study of Self-Peer, Peer-Peer, and Self-Self (Dis)Agreement. *Journal of Personality, 78*(5), 1565-1594. doi:10.1111/j.1467-6494.2010.00661.x

Kandler, C., Zimmermann, J., & McAdams, D. P. (2014). Core and Surface Characteristics for the Description and Theory of Personality Differences and Development. *European Journal of Personality, 28*(3), 231-243. doi:10.1002/per.1952

Kenny, D. A. (2004). PERSON: A general model of interpersonal perception. *Personality and Social Psychology Review, 8*(3), 265-280. Retrieved from <Go to ISI>://000223500400003

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the Big Five. *Psychological Bulletin, 116*(2), 245-258. Retrieved from <Go to ISI>://A1994PF97900003

Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. C. (1992). Consensus at zero acquaintance: replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62*(1), 88-97.

Kenny, D. A., & Kashy, D. A. (1994). Enhanced co-orientation in the perception of friends: a social relations analysis. *Journal of Personality and Social Psychology, 67*(6), 1024-1033.

Kenny, D. A., & West, T. V. (2010). Similarity and Agreement in Self-and Other Perception: A Meta-Analysis. *Personality and Social Psychology Review, 14*(2), 196-213. doi:10.1177/1088868309353414

Kenny, D. A., & Winquist, L. A. (2001). The measurement of interpersonal sensitivity: Consideration of design, components, and unit of analysis In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 265-302). Englewood Cliffs, NJ: Erlbaum.

Klein, K. J. K., & Hedges, S. D. (2001). Gender differences, motivation, and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin, 27*(6), 720-730. doi:10.1177/0146167201276007

Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality, 64*, 311-337.

Kraus, M. W., Cote, S., & Keltner, D. (2010). Social Class, Contextualism, and Empathic Accuracy. *Psychological Science, 21*(11), 1716-1723. doi:10.1177/0956797610387613

Kurtz, J. E., & Sherker, J. L. (2003). Relationship quality, trait similarity, and self-other agreement on personality ratings in college roommates. *Journal of Personality, 71*(1), 21-48. Retrieved from 12597236

Laurent, S. M., & Hodges, S. D. (2009). Gender Roles and Empathic Accuracy: The Role of Communion in Reading Minds. *Sex Roles, 60*(5-6), 387-398. doi:10.1007/s11199-008-9544-x

Lee, K., & Ashton, M. C. (2008). The HEXACO Personality Factors in the Indigenous Personality Lexicons of English and 11 Other Languages. *Journal of Personality, 76*(5), 1001-1053. doi:10.1111/j.1467-6494.2008.00512.x

Lee, K., Ashton, M. C., Pozzebon, J. A., Visser, B. A., Bourdage, J. S., & Ogunfowora, B. (2009). Similarity and Assumed Similarity in Personality Reports of Well-Acquainted Persons. *Journal of Personality and Social Psychology, 96*(2), 460-472. doi:10.1037/a0014059

Leikas, S., Verkasalo, M., & Lönnqvist, J.-E. (2013). Posing personality: Is it possible to enact the Big Five traits in photographs? *Journal of Research in Personality, 47*(1), 15-21. doi:10.1016/j.jrp.2012.10.012

Leitsalu, L., Haller, T., Esko, T., Tammesoo, M.-L., Alavere, H., Snieder, H., . . . Metspalu, A. (2014). Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *International Journal of Epidemiology, 44*. doi:doi: 10.1093/ije/dyt268

Letzring, T. D. (2008). The good judge of personality: Characteristics, behaviors, and observer accuracy. *Journal of Research in Personality, 42*(4), 914-932. doi:10.1016/j.jrp.2007.12.003

Letzring, T. D., & Human, L. J. (2014). An Examination of Information Quality as a Moderator of Accurate Personality Judgment. *Journal of Personality, 82*(5), 440-451. doi:10.1111/jopy.12075

Letzring, T. D., Wells, S. A., & Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology, 91*(1), 111-123. Retrieved from <Go to ISI>://000239139900008

Levesque, M. J., & Kenny, D. A. (1993). Accuracy of behavioral predictions at zero acquaintance: A social-relations analysis. *Journal of Personality and Social Psychology, 65*(6), 1178-1187. Retrieved from <Go to ISI>://A1993MM03900008

Lippa, R. A. (2010). Sex Differences in Personality Traits and Gender-Related Occupational Preferences across 53 Nations: Testing Evolutionary and Social-Environmental Theories. *Archives of Sexual Behavior, 39*(3), 619-636. doi:10.1007/s10508-008-9380-7

Luo, S. H., & Klohnen, E. C. (2005). Assortative mating and marital quality in newlyweds: A couple-centered approach. *Journal of Personality and Social Psychology, 88*(2), 304-326. Retrieved from <Go to ISI>://000226585400006

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality, 63*, 365-396.

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *Plos One, 9*(3). doi:10.1371/journal.pone.0090779

McCrae, R. R. (2008). A note on some measures of profile agreement. *Journal of Personality Assessment, 90*, 105-109.

McCrae, R. R., & Costa, P. T. (1982). Self-concept and the stability of personality: Cross-sectional comparisons of self-reports and ratings. *Journal of Personality and Social Psychology, 43*(6), 1282-1292.

McCrae, R. R., Costa, P. T., & Martin, T. A. (2005). The NEO-PI-3: A More Readable Revised Neo Personality Inventory. *Journal of Personality Assessment, 84*(3), 261-270. Retrieved from ISI:000229342900005

McCrae, R. R., Costa, P. T., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., . . . Urbanek, T. (2004). Consensual Validation of Personality Traits Across Cultures. *Journal of Research in Personality, 38*(2), 179-201. Retrieved from ISI:000189245900006

McCrae, R. R., Martin, T. A., Hrebickova, M., Urbanek, T., Boomsma, D. I., Willemsen, G., & Costa, P. T. (2008). Personality trait similarity between spouses in four cultures. *Journal of Personality, 76*(5), 1137-1163. doi:10.1111/j.1467-6494.2008.00517.x

McLarney-Vesotski, A., Bernieri, F., & Rempala, D. (2011). An experimental examination of the "good judge". *Journal of Research in Personality, 45*(4), 398-400. doi:10.1016/j.jrp.2011.04.005

Mõttus, R., Allik, J., & Pullman, H. (2007). Does personality vary across ability levels? A study using self and other ratings. *Journal of Research in Personality, 41*(1), 155-170. Retrieved from <Go to ISI>://000244840700011

Mõttus, R., Realo, A., Allik, J., Esko, T., Metspalu, A., & Johnson, W. (2015). Within-Trait Heterogeneity in Age Group Differences in Personality Domains and Facets: Implications for the Development and Coherence of Personality Traits. *Plos One, 10*(3). doi:10.1371/journal.pone.0119667

Murphy, N. A., & Hall, J. A. (2011). Intelligence and interpersonal sensitivity: A meta-analysis. *Intelligence, 39*(1), 54-63. doi:10.1016/j.intell.2010.10.001

Obschonka, M., Schmitt-Rodermund, E., Silbereisen, R. K., Gosling, S. D., & Potter, J. (2013). The Regional Distribution and Correlates of an Entrepreneurship-Prone Personality Profile in the United States, Germany, and the United Kingdom: A Socioecological Perspective. *Journal of Personality and Social Psychology, 105*(1), 104-122. doi:10.1037/a0032275

Paunonen, S. V. (1989). Consensus in personality judgements: moderating effects of target-rater acquaintenceship and behavior observability. *Journal of Personality and Social Psychology, 56*, 823-833.

Paunonen, S. V., & Kam, C. (2014). The accuracy of roommate ratings of behaviors versus beliefs. *Journal of Research in Personality, 52*, 55-67. doi:10.1016/j.jrp.2014.07.006

Pelham, B. W. (1993). The idiographic nature of human personality: examples of the idiographic self-concept. *Journal of Personality and Social Psychology, 64*(4), 665-677.

Realo, A., Allik, J., Nõlvak, A., Valk, R., Ruus, T., Schmidt, M., & Eilola, T. (2003). Mind-reading ability: Beliefs and performance. *Journal of Research in Personality, 37*(5), 420-445. Retrieved from <Go to ISI>://000185363500005

Realo, A., Allik, J., & Vadi, M. (1997). The hierarchical structure of collectivism. *Journal of Research in Personality, 31*, 93-116.

Robbins, J. M., & Krueger, J. I. (2005). Social Projection to Ingroups and Outgroups: A Review and Meta-Analysis. *Personality and Social Psychology Review, 9*(1), 32-47. doi:10.1207/s15327957pspr0901_3

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of Mean-Level Change in Personality Traits Across the Life Course: a Meta-Analysis of Longitudinal Studies. *Psychological Bulletin, 132*(1), 1-25. doi:10.1037/0033-2909.132.1.1

Roberts, B. W., Wood, D., & Smith, J. L. (2005). Evaluating Five Factor Theory and Social Investment Perspectives on Personality Trait Development. *Journal of Research in Personality, 39*(1), 166-184. Retrieved from ISI:000226267200010

Rogers, K. H., & Biesanz, J. C. (2015). Knowing Versus Liking: Separating Normative Knowledge From Social Desirability in First Impressions of Personality. *Journal of Personality and Social Psychology, 109*(6), 1105-1116. doi:10.1037/a0039587

Runeson, S. (1977). On the possibility of "smart" perceptual mechanisms. *Scandinavian Journal of Psychology, 18*, 172-179.

Schneider, D. J., Hastorf, A. H., & Ellsworth, P. (1979). *Person perception* (2d ed.). Reading, Mass.: Addison-Wesley Pub. Co.

Schneider, L., Schimmack, U., Petrican, R., & Walker, S. (2010). Acquaintanceship length as a moderator of self-informant agreement in life-satisfaction ratings. *Journal of Research in Personality, 44*(1), 146-150. doi:http://dx.doi.org/10.1016/j.jrp.2009.11.004

Shevlin, M., Walker, S., Davies, M. N. O., Banyard, P., & Lewis, C. A. (2003). Can you judge a book by its cover? Evidence of self-stranger agreement on personality at zero acquaintance. *Personality and Individual Differences, 35*(6), 1373-1383. doi:10.1016/s0191-8869(02)00356-2

Sorokowska, A., Sorokowski, P., & Szmajke, A. (2012). Does Personality Smell? Accuracy of Personality Assessments Based on Body Odour. *European Journal of Personality, 26*(5), 496-503. doi:10.1002/per.848

South, S. C., Oltmanns, T. F., Johnson, J., & Turkheimer, E. (2011). Level of Agreement Between Self and Spouse in the Assessment of Personality Pathology. *Assessment, 18*(2), 217-226. doi:10.1177/1073191110394772

Starzyk, K. B., Holden, R. R., Fabrigar, L. R., & Macdonald, T. K. (2006). The Personal Acquaintance Measure: a Tool for Appraising One's Acquaintance With Any Person. *Journal of Personality and Social Psychology, 90*(5), 833-847. Retrieved from ISI:000238083400008

Story, A. L. (2003). Similarity of trait construal and consensus in interpersonal perception. *Journal of Experimental Social Psychology, 39*(4), 364-370. doi:10.1016/s0022-1031(02)00529-2

Swann, W. B., & Gill, M. J. (1997). Confidence and accuracy in person perception: Do we know what we think we know about our relationship partners? *Journal of Personality and Social Psychology, 73*(4), 747-757. Retrieved from <Go to ISI>://A1997XX99100008

Swider, B. W., Zimmerman, R. D., Charlier, S. D., & Pierotti, A. J. (2015). Deep-level and surface-level individual differences and applicant attraction to organizations: A meta-analysis. *Journal of Vocational Behavior, 88*, 73-83. doi:10.1016/j.jvb.2015.01.005

Tackett, J. L., Herzhoff, K., Kushner, S. C., & Rule, N. (2016). Thin Slices of Child Personality: Perceptual, Situational, and Behavioral Contributions. *Journal of Personality and Social Psychology, 110*(1), 150-166. doi:10.1037/pspp0000044

Taft, R. (1955). The ability to judge people. *Psychological Bulletin, 52*(1), 1-23. doi: 10.1037/h0044999

Thorndike, R. L. (1949). *Personnel selection; test and measurement techniques*. New York,: J. Wiley.

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review, 41*, 1-32.

Vazire, S. (2010). Who Knows What About a Person? The Self-Other Knowledge Asymmetry (SOKA) Model. *Journal of Personality and Social Psychology, 98*(2), 281-300. doi:10.1037/a0017908

Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: the role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology, 78*(3), 546-558. Retrieved from 10743880

Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: the affordances of a Gibsonian approach. *Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc, 1*(3), 204-223. doi:10.1207/s15327957pspr0103_2

Figure Captions

Figure 1. Self-other trait correlations for the 30 subscales of NEO PI-3 (A) and HEXACO-PI-R (B).

Figure 2. Distributions of the self-other profile correlations for NEO PI-3 and HEXACO-PI-R.

Figure 3. Distinctive self-other profile agreement dependent on target age in the Estonian (A) and Dutch (B) samples.

Figure 4. The mean profile correlation for possible target-judge combinations.

Figure 5. The mean profile correlation for the three categories of relationship between judges and targets.

Figure 1

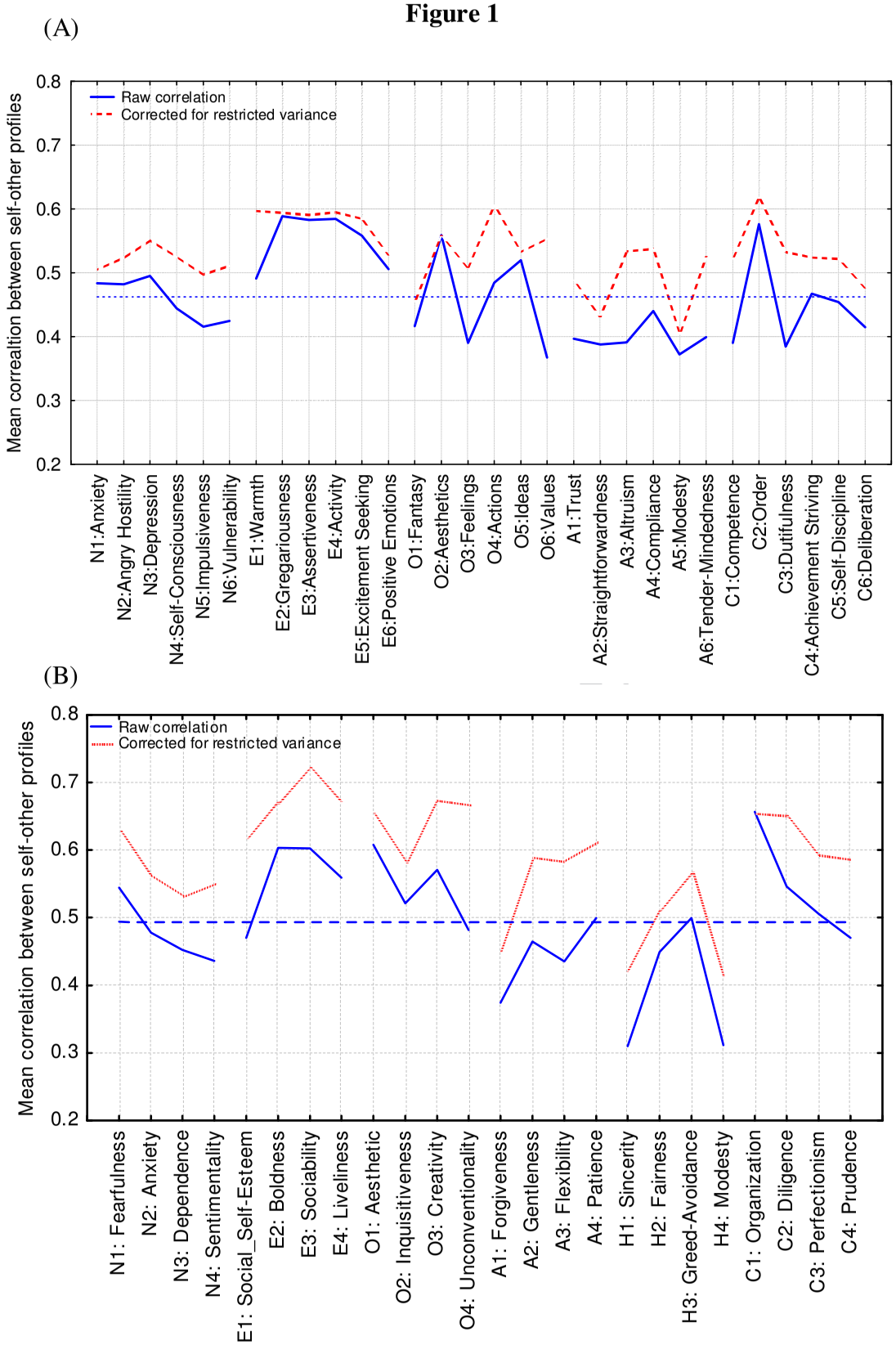


Figure 2

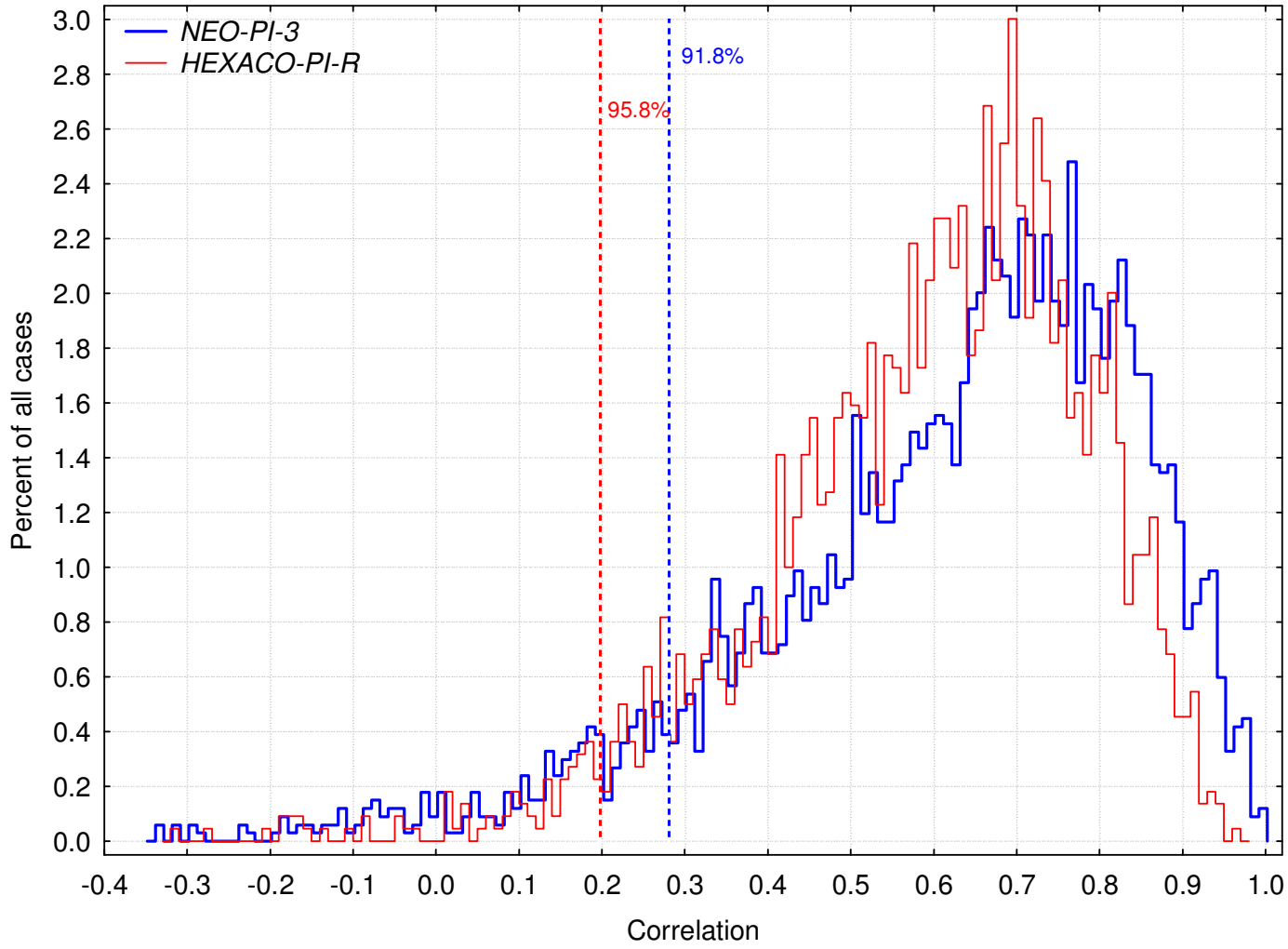


Figure 3

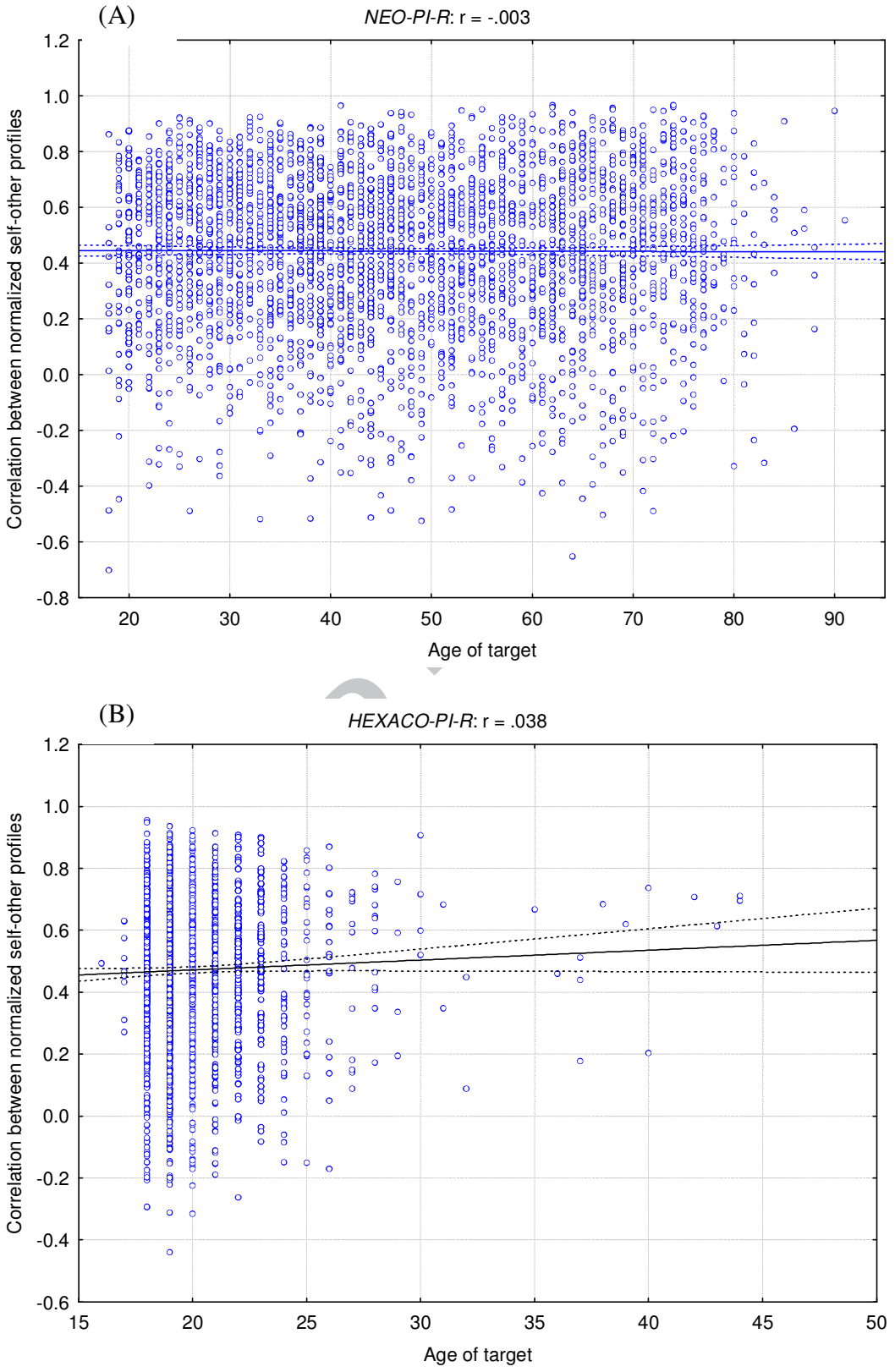


Figure 4

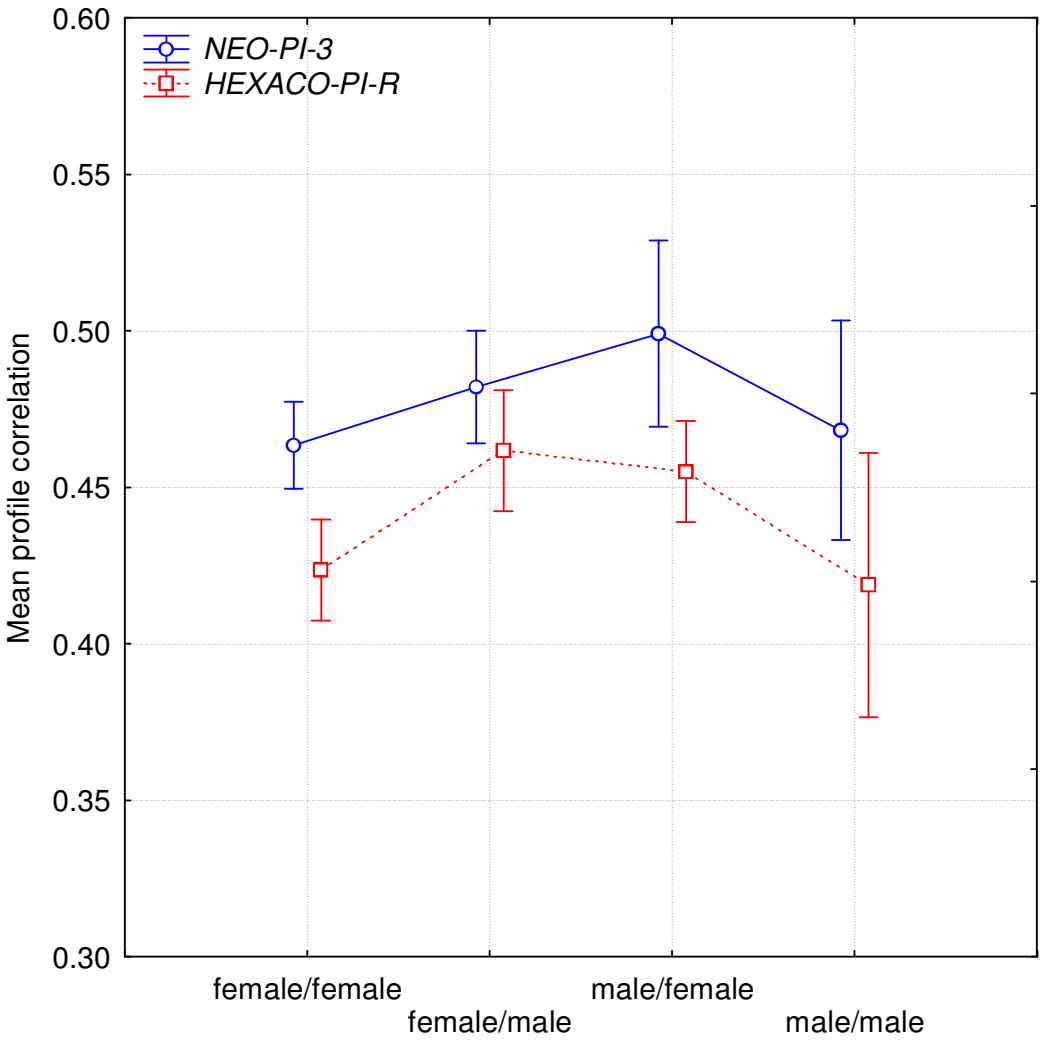
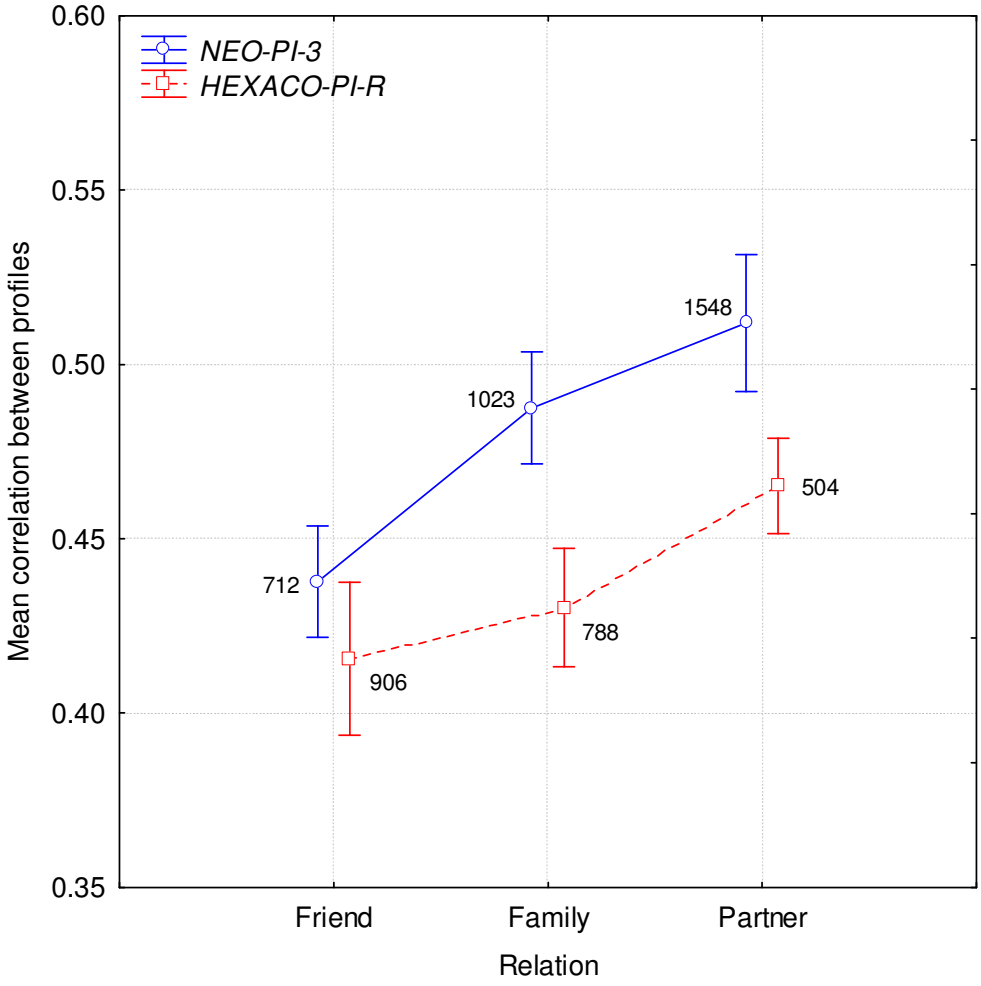


Figure 5



- Two large samples (N = 11,096) completed two different personality questionnaires, the *NEO-PI-3* and *HEXACO-PI-R*.
- Age and sex of participants and length of acquaintance had no significant effect on the level of self-other agreement.
- Sufficiently good agreement between self- and other-judgments was reached in most target-judge pairs.
- Because judging another person's personality is an easy task, it is difficult to determine "good targets," "good judges," or "good traits."
- Perhaps only the closeness of the target-judge relationship moderates the level of self-other agreement.