Journal of Personality and Social Psychology
2010, Vol. 99, No. 5, 870–882

© 2010 American Psychological Association
0022-3514/10/$12.00 DOI: 10.1037/a0020963

How People See Others Is Different From How People See Themselves: A Replicable Pattern Across Cultures

Jüri Allik, Anu Realo, and René Mõttus
University of Tartu

Peter Borkenau
Martin-Luther-Universität Halle–Wittenberg

Peter Kuppens
University of Melbourne and Katholieke Universiteit Leuven

Martina Hřebíčková
Academy of Sciences of the Czech Republic

Consensus studies from 4 cultures—in Belgium, the Czech Republic, Estonia, and Germany—as well as secondary analyses of self- and observer-reported Revised NEO Personality Inventory (NEO PI-R) data from 29 cultures suggest that there is a cross-culturally replicable pattern of difference between internal and external perspectives for the Big Five personality traits. People see themselves as more neurotic and open to experience compared to how they are seen by other people. External observers generally hold a higher opinion of an individual's conscientiousness than he or she does about him- or herself. As a rule, people think that they have more positive emotions and excitement seeking but much less assertiveness than it seems from the vantage point of an external observer. This cross-culturally replicable disparity between internal and external perspectives was not consistent with predictions based on the actor–observer hypothesis because the size of the disparity was unrelated to the visibility of personality traits. A relatively strong negative correlation ($r = -.53$) between the average self-minus-observer profile and social desirability ratings suggests that people in most studied cultures view themselves less favorably than they are perceived by others.

Keywords: personality ratings, internal and external perspective, cross-cultural comparison, self-enhancement, the actor–observer hypothesis

More than a decade ago, McCrae and Costa (1997) proposed the hypothesis that the pattern of covariation among basic personality traits is a universal feature of the human species. Several recent large-scale cross-cultural studies have supported this idea, showing that a common five-factor structure of personality traits can be found in all languages and cultures examined so far (McCrae, Terracciano, & 78 Members of the Personality Profiles of Cultures Project, 2005; Schmitt, Allik, McCrae, & Benet-Martinez, 2007). However, the five-factor model is not the only replicable personality structure because Eysenck's three-factor (van Hemert, van de Vijver, Poortinga & Georgas, 2002) and psycholexical six-factor (Lee & Ashton, 2008) structures have also been replicated in many cultures.

Jüri Allik, Anu Realo, and René Mõttus, Department of Psychology, University of Tartu, Tartu, Estonia; Peter Borkenau, Institut für Psychologie, Martin-Luther-Universität Halle–Wittenberg, Halle, Germany; Peter Kuppens, School of Behavioral Science, University of Melbourne, Melbourne, Australia, and Department of Psychology, Katholieke Universiteit Leuven, Leuven, Belgium; Martina Hřebíčková, Institute of Psychology, Academy of Sciences of the Czech Republic, Prague, the Czech Republic.

This project was supported by Estonian Ministry of Science and Education Grant SF0180029s08 and Estonian Science Foundation Grant ESF7020 to Jüri Allik. The Czech participation was supported by Grant Agency of the Czech Republic Grant P407/10/2394. We are very thankful to Robert R. McCrae for his helpful comments and suggestions.

Correspondence concerning this article should be addressed to Jüri Allik, Department of Psychology, University of Tartu, Tiigi 78, Tartu 50410, Estonia. E-mail: juri.allik@ut.ee

In addition to the consistent pattern of covariation among personality traits, several other surprisingly universal features of personality have been discovered. For example, it was found that in almost every culture, women reported themselves to be higher in Neuroticism, Agreeableness, Warmth, and Openness to Feelings, whereas men were higher in Assertiveness and Openness to Ideas (Costa, Terracciano, & McCrae, 2001; Feingold, 1994; Schmitt, Realo, Voracek, & Allik, 2008). Quite unexpectedly, these sex differences in personality increased with higher levels of human development, including a long and healthy life, equal access to knowledge and education, and economic prosperity (Costa et al., 2001; Schmitt et al., 2008). What is particularly remarkable is that the cross-cultural convergence between different studies on sex differences in personality was demonstrably stronger and more replicable than the convergence between the mean levels of the traits themselves (Schmitt et al., 2007, 2008).

Another feature that seems to easily transcend cultures is age difference. In every human society explored so far, individuals become less extraverted and open to new experiences and more agreeable and conscientious with age (Costa et al., 2000; McCrae, Costa, Hřebíčková et al., 2004; McCrae et al., 1999; Srivastava, John, Gosling, & Potter, 2003). In spite of some specific features characterizing particular cultures, the general pattern of difference between younger and older individuals is stable and highly replicable across the 30 personality traits measured by the NEO PI-R questionnaire, even across dissimilar cultures (Allik et al., 2009).

In light of the discovery of these universal features, it is surprising that there is no consensus about the systematic differences

between how people see others and how they see themselves, despite some theories that have been proposed. Social psychologists, for example, have invested a considerable amount of energy in the promotion of the idea that there is a fundamental disparity between the way people perceive themselves and the way they are perceived by others (Jones & Nisbett, 1971; Nisbett, Caputo, Legant, & Marecek, 1973; Watson, 1982). This disparity is believed to originate from an inevitable asymmetry between internal and external viewpoints: People are immersed in their own sensations, emotions, and cognitions at the same time that their experience of others is dominated by what can be observed externally (Pronin, 2008). Our aim in the current article is to systematically analyze the differences between how people judge their personality and how their personality is judged by others across different cultures. Analyzing the existing literature, it is possible to distinguish at least two principal mechanisms by which personality descriptions made from the perspective of the first person might be systematically different from those made from the vantage point of the third person.

Different Information

As an external observer cannot see the target person in all situations, it is almost inevitable that the information possessed by the target person must be different from the information that is available to an external observer. Even when it comes to information that is equally available to the target person and the external observer, it is possible that the target person could ignore some information that was attended to by the external observer in making his or her judgment. One consequence of this, as argued by Jones and Nisbett (1971), is the actor–observer asymmetry in attribution, that is the pervasive tendency for actors to attribute their behavior to external (situational) causes and for observers to attribute the same behavior to internal causes (dispositional qualities) of the actor. As a further evidence of the divergent perspectives of the actor and observer, among other things observers were found to ascribe more personality traits to other people than to themselves (Nisbett et al., 1973). However, a recent meta-analysis involving more than 170 studies established that the actor–observer hypothesis—counter to what textbook descriptions and commonly held beliefs suggest—is neither firmly established nor robust and that evidence for it is surprisingly limited (Malle, 2006). For example, the established beliefs that people ascribe more personality traits to others than to themselves (Sande, Goethals, & Radloff, 1988) and that they perceive more complexity in their behavior than that of others (Locke, 2002) were not confirmed in these later studies.

It is not self-evident how the principle of "more personality traits" can be operationalized for responses to a fixed list of items. Personality psychology questionnaires are constructed on the assumption that all personality traits are applicable to all individuals; respondents thus need to assess themselves or the others on each of these traits. Despite the fact that considerable self–other agreement can be obtained for all personality traits, the convergence of judgments either between self and observer or between different observers is considerably stronger on some traits than on the other traits (Colvin, 1993; Funder, 1999; John & Robins, 1993). The most likely cause for these differences is the type and amount of information available to the self and the external observer. In particular, results suggest that traits pertaining to extraversion are revealed relatively directly in social behavior and, therefore, are easy to judge, whereas traits pertaining to neuroticism are less visible and, so, are judged less accurately (Funder & Colvin, 1988; Funder & Dobroth, 1987). It has also been noticed that traits reflecting affective states are more difficult to judge than are traits which manifest themselves in overt behaviors (Spain, Eaton, & Funder, 2000; Watson, Hubbard, & Wiese, 2000). Lack of judgeability or visibility of traits can lead, in principle at least, to a disparity between self- and observer-ratings (John & Robins, 1993). It is not difficult to imagine, for example, why people see themselves as more neurotic compared to how they are seen by other people. Neuroticism reflects, to a considerable extent, inner states of an individual that are not necessarily accessible to an external observer. Provided that these externally unobservable instances of neurotic tendencies influence people's self-evaluation of neuroticism, the expected outcome is a disparity between self- and observer-ratings. Thus, one possible scenario is that the self–other disparity is less pronounced on more observable traits—as operationalized using rank-order correlations between self- and observer-ratings—whereas the disparity between rater's perspectives is more manifest in less externally observable traits.

Self-Enhancement

One of the most pervasive explanations for the disparity between self- and observer-perceptions is that people are systematically engaged in self-enhancement: They view themselves more favorably than they view others (Kenny, 1994; Kwan, John, Kenny, Bond, & Robins, 2004). Although some cross-cultural differences seem to exist (Heine & Hamamura, 2007; Heine & Renshaw, 2002), a recent cross-cultural study demonstrated that in all 56 cultures studied, people's mean value of self-esteem was above the scale neutral point (Schmitt & Allik, 2005), suggesting that most people are motivated to maintain a positive view of themselves. Many studies have demonstrated that college students rate their own personality traits in more socially desirable terms than they do when rating the "average college student" (Alicke, 1985; Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995; J. Krueger, 1998). However, the effect of this unrealistically positive view of themselves disappears, or is considerably reduced, when a specific person, not an average college student, is assessed (Alicke et al., 1995). Nevertheless, it is possible that the mean difference between self- and observer-ratings reflects, to a certain degree, the social desirability of personality traits. Thus, one expected outcome of self-enhancement is that mean self-rating scores are higher than are the observer-ratings on those personality traits believed to have higher social desirability.

However, all these predictions about systematic differences between how people see others' and their own personality traits still need to take into account the fact that normative self-rated personality mean scores converge almost perfectly with normative observer-rated mean scores. For example, the correlation between the mean profiles of the adult S-Form (self-ratings) and R-Form (observer-ratings) presented in the NEO PI-R Professional Manual is .94 (Costa & McCrae, 1992, Table B1 and Table B2). The largest mean difference is 3.3 raw-score points on the C6: Deliberation subscale, which in *T*-scores is 7.6 units higher for observer- than self-ratings. The average absolute self–other rating difference

across all 30 NEO PI-R subscales is 2.9 *T*-score units. Thus, given the high correlation and small difference in mean levels, there is little room for disparity between the perspectives from which personality is described. Looking at how similar self-rating and observer-rating normative profiles are, one has to conclude that the effect caused by self–other asymmetry must be small and, if at all, reliably detectable only when sufficiently large samples are used.

Examination of the global distribution of aggregated self-rated personality profiles across cultures has revealed a regular pattern, with a clear contrast between European and American cultures on the one hand and Asian and African cultures on the other (Allik & McCrae, 2004). The global distribution of observer-rated personality traits generally followed the same pattern (McCrae, Terracciano, & 79 Members of the Personality Profiles of Cultures Project, 2005). However, it is possible that when there is an asymmetry between self- and observer-perspectives, it also displays itself at the aggregate level. If there is a systematic disparity in the perspective from which personality traits are described at culture-level personality scores as well, it will constitute another cross-cultural personality universal in addition to the relatively invariant pattern of sex and age differences.

Aims of the Study

Since many theoretical constructions assume a fundamental asymmetry between self- and observer-perspectives, the main goal of this study is to examine whether there is a replicable pattern of differences between self-rated and observer-rated personality traits which transcends different languages and cultures. There are two principal study designs available: an individual design, in which the targets of self-reports and observer reports are the same and in which the question is whether the mean difference between these two is positive or negative; and a culture-level design, in which different samples from the same culture are compared. The former is more powerful because targets serve as their own controls with regard to trait level, but the latter is also informative since it helps to test the robustness of the phenomenon. We used both designs.

In Study 1 we compared data from four European languages and cultures—in Belgium, Czech Republic, Estonia, and Germany—where personality traits of participants were judged by themselves and by one or more observers. To many readers it may come as a surprise that the number of studies with consensual validation between self- and observer-ratings is rather limited (McCrae, Costa, Martin et al., 2004). Even so, consensus studies are the most powerful source to test whether how people see others is different from how people see themselves. Following the idea of the actor–observer hypothesis, we tested whether self–observer discrepancy is larger in less visible traits. Additionally, separate groups of judges from all four of the countries also rated the social desirability of each of the NEO-PI-R items. On the basis of this, it is possible to compare the difference between self- and observer-ratings with the social desirability of personality traits. Self-rated scores are expected to be higher than observer-rated scores on those personality traits considered higher on social desirability.

Study 2 is devoted to analyses of published cross-cultural data sets about self- and observer-ratings of personality with the NEO PI-R. The main goal in Study 2 is to establish how robust are the asymmetries between self- and observer-ratings that were found in the first study, that is how well they generalize across languages, cultures, and different levels of analyses.

Study 1

Samples

Belgian sample. Flemish data were collected from 345 target participants (270 women and 75 men) who were psychology students at the Katholieke Universiteit Leuven and who, as a course requirement, rated their own personality with the Dutch version of the NEO PI-R (Hoekstra, Ormel, & DeFruyt, 1996). They also recruited a well-acquainted person ($n = 345$; 190 women, 112 men, and 43 did not specify sex), either a relative or a friend, who rated their personality using the observer-report form of the same instrument. The mean age of targets was 18.4 ($SD = 3.0$) years. The mean age of external raters was 29.5 ($SD = 13.7$) years.

Czech sample. The Czech sample included 811 targets (330 men, 481 women) who were recruited in a series of studies (McCrae, Costa, Martin et al., 2004). They ranged in age from 14 years to 83 years, with a mean age of 35.7 years ($SD = 14.2$ years). Peer-ratings were provided by 909 raters (377 men, 532 women) aged 14–83 years ($M = 35.8$ years; $SD = 14.3$ years) who participated in one of two research designs. In the self–other agreement studies ($N = 615$), each target provided a self-report and was rated by one informant. In the consensus study, 195 targets (85 men and 110 women aged 17–77 years; mean age 36.4, $SD = 15.2$) provided a self-report and were each rated by three informants. All participants used the Czech version of the NEO PI-R questionnaire (Hřebíčková, 2002).

Estonian sample. Estonian data came from two already published studies. The first subsample consisted of 218 Estonian-speaking participants (180 women and 38 men; mean age 22.3 years, $SD = 5.2$) who answered the NEO PI-R questionnaire, which was accompanied by a standard instruction to describe themselves honestly and accurately (Konstabel, Aavik, & Allik, 2006). They were also asked to provide two peer-reports ($n = 436$) from their acquaintances, relatives, or close friends. The Estonian version of the NEO PI-R (Kallasmaa, Allik, Realo, & McCrae, 2000) was completed voluntarily; some students studying psychology received an extra credit towards the fulfillment of their course requirements. The second Estonian subsample consisted of 154 participants (53 men and 101 women; mean age 43.9 years, $SD = 17.6$) who were described by one or two judges (Mõttus, Allik, & Pullman, 2007). The sample of judges ($n = 308$) included 203 women, 67 men, and 38 participants who did not report their gender. The mean age of the judges was 38.2 ($SD = 15.9$) years. Both targets and judges used the Estonian version of the EE.PIP-NEO (Mõttus, Pullmann, & Allik, 2006), which has a facet-structure identical to the NEO PI-R but was designed to be linguistically simpler, containing shorter and grammatically less-complex items.

German sample. Participants were 304 students (169 women, 134 men, 1 not reporting sex) at a German university, of whom only 3 studied psychology (Borkenau & Zaltauskas, 2009). Their mean age was 23.38 ($SD = 2.68$) years, ranging from 18 years to 35 years. They received 45 Euro for their participation and were recruited in 76 groups, each comprising 4 persons who all knew

each other well. First, the participants described the 3 other group members on 30 bipolar adjective scales; these, however, are not relevant to the present study. Next, each 4-person group was split into two dyads, and all participants described themselves and the other dyad member on several personality inventories including the German version of the NEO PI-R (Ostendorf & Angleitner, 2004). It is important to notice that in all four samples observers knew their targets well, being either their close relatives or friends.

Measure of Social Desirability

In order to develop a social desirability index for the NEO PI-R, the questionnaire items were assessed by 100 Czech judges (43 men and 57 women, mean age 40.5 years, $SD = 15.1$), 88 Estonian judges (24 men and 64 women, mean age 37.6 years, $SD = 12.7$), 30 Flemish judges (12 men, 16 women, 2 did not report their sex, mean age 20.3 years, $SD = 2.0$), and 20 German judges (9 men and 11 women, mean age 23.8 years, $SD = 3.0$), who independently rated the social desirability of each of the 240 NEO PI-R items. The instruction stated,

> Descriptors of people often contain evaluative information. Some personality characteristics are considered more desirable, receiving approval from other people, whereas others are undesirable. If someone agrees strongly with this item, does this present that person in a favorable or unfavorable light, or is agreeing with this item neutral as regards to others' approval?

Ratings were made on a 7-point Likert scale, ranging from extremely undesirable (−3) to extremely desirable (+3), with zero as a neutral point. Estonian desirability ratings were reported in a previous study (Konstabel et al., 2006). As the pairwise correlations between social desirability profiles of cultures were sufficiently high (from $r = .80$ to $r = .93$), we used an unweighted average of these four groups of judges.

Results and Discussion

The Flemish, German, Estonian, and Czech mean difference profiles (self-minus-observer) converted to z -scores (the mean difference divided by the average standard deviation) are shown in Figure 1. Even a brief inspection reveals that all four difference profiles are very similar. Indeed, the pairwise correlations between these four profiles range from .77 (Belgium and Germany) to .88 (Estonia and the Czech Republic) with the mean $r = .83$ (all highly significant). All four profiles are also strongly correlated with the difference profile of the U.S. adult normative data (S-Form minus R-Form) given in the NEO PI-R Professional Manual (Costa & McCrae, 1992, Table B1 and Table B2). The correlations are .85, .79, .75, and .63 for Belgium, the Czech Republic, Estonia, and Germany, respectively.

In all four cultures (in addition to the United States), people perceive themselves as more neurotic and more open than they are seen by others. They also perceive their level of conscientiousness as lower than how they are seen from the vantage point of their observers. Particularly, people see themselves as less competent (C1), self-disciplined (C5), and altruistic (A3) than they are perceived by others. At the same time they think that they are more than others open to fantasy (O1) and ready to reexamine social, political, and religious values (O6). Even this short list of disparities suggests that in general, people do not view themselves more favorably than how they are viewed by others.

For a more formal test we analyzed the combined sample of 1,768 targets, pooled from the four separate samples. Before pooling, all scores were normalized within the country dataset.

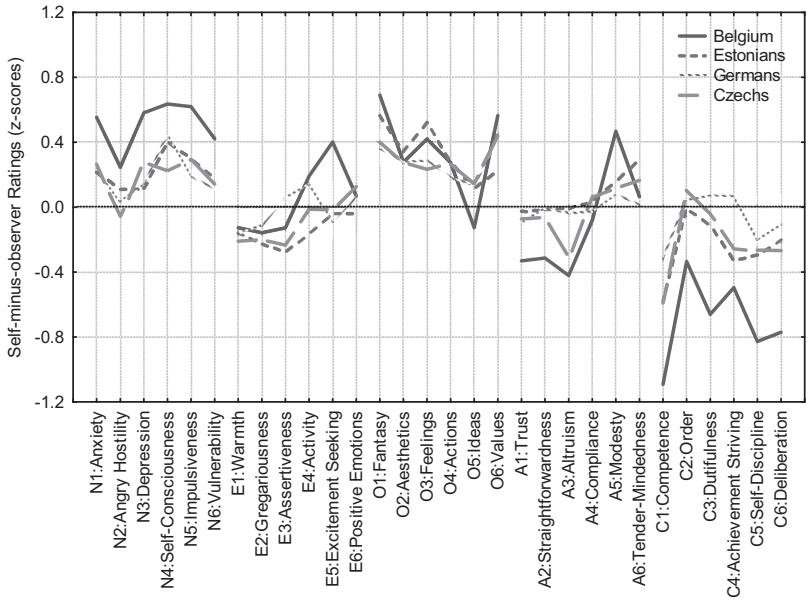

Figure 1. The mean difference profiles (self-minus-observer) for Belgium, Germany, Estonia, and the Czech Republic, converted to z-scores (the mean difference divided by the average standard deviation). Letters–number combinations are the NEO PI-R facet scale numbers.

Contrary to the concept of self-enhancement, the mean self-minus-observer profile was negatively correlated ($r = -.53$, $p = .003$), with the profile of social desirability ratings (individual correlations were from $r = -.40$, $p = .026$ to $r = -.62$, $p = .001$ for Estonia and Belgium, respectively) suggesting that observer-ratings rather than self-ratings might be biased towards social desirability. Similarly, the normative U.S. adult self-minus-observer profile was negatively correlated with the profile of social desirability ratings ($r = -.45$, $p = .012$). Figure 2 presents the average self-minus-observer profile for the four cultures studied in comparison with their average social desirability ratings on the items of the 30 NEO PI-R facet scales. Although these two profiles are not exact mirror images, their dissimilarity is obvious. In personality psychology, a correlation of −.53 between two profiles is sufficiently high enough to speak of a substantial reverse link between social desirability and the disparity between perspectives.

To test the prediction following from the actor–observer asymmetry that disparity is more pronounced on less visible traits, we computed the rank-order correlation between self-ratings and observer-ratings. The highest self–other agreement was found on the subscales E3: Assertiveness (.56), O2: Aesthetics (.56), and E5: Excitement Seeking (.55). A relatively low self–other agreement was found in the ratings of C1: Competence (.32), A1: Trust (.36), and N5: Impulsiveness (.38). The average self–other agreement was reasonably high (median = .43, $p = .018$). These values are comparable to typical self–observer agreement values obtained in previous studies using the NEO family questionnaires (Connolly, Kavanagh, & Viswesvaran, 2007; McCrae, Costa, Martin et al., 2004). Having obtained these self–other convergence values, we can ask whether the self–observer asymmetry is more pronounced on those personality traits on which people and their judges agree less. Figure 3 shows the correlation plot between self–other agreement and self-minus-observer difference scores. There seems to be no systematic relationship between self–other agreement and the asymmetry in perception ($r = .03$, $p = .88$). Thus, on personality traits that are more visible or judge-able, there is no smaller disparity between self- and observer-ratings.

In addition to social desirability and visibility, there are other potential explanations for the asymmetry between self- and observer-ratings. One of them could be a systematic age difference between targets and judges. College-age targets may have recruited adults to judge their personality, and adults may tend to endorse certain items differently compared with college-aged people. Indeed, this tendency was true for the Flemish and Estonian samples, but not for the German and a part of Czech samples in which the mean age of target and judges was identical. Thus, if the rater-age related explanation is true, we should expect a systematic difference in self–observer asymmetry between German and Czech samples on one side and Estonian and Flemish samples on the other side. There appears to be no such difference (see Figure 1), which speaks against the rater-age related explanation for the asymmetry. In addition, to see whether the observer-rated personality profile tends to be more adultlike, we compared the average values of the four self-minus-observer profiles displayed in Figure 1 with the mean age differences between adults and college-age targets for 30 NEO PI-R facets. We took the latter values from the best observer-ratings database to date, results obtained from the international sample consisting of 50 cultures participating in the Personality Profiles of Cultures Project (McCrae, Terracciano, & 78 Members of the Personality Profiles of Cultures Project, 2005; McCrae, Terracciano, & 79 Members of the Personality Profiles of Cultures Project, 2005). Contrary to our prediction that adults may follow their age-specific response pattern even when rating college-aged people, the correlation between the two difference profiles— self–observer asymmetry and age-related differences

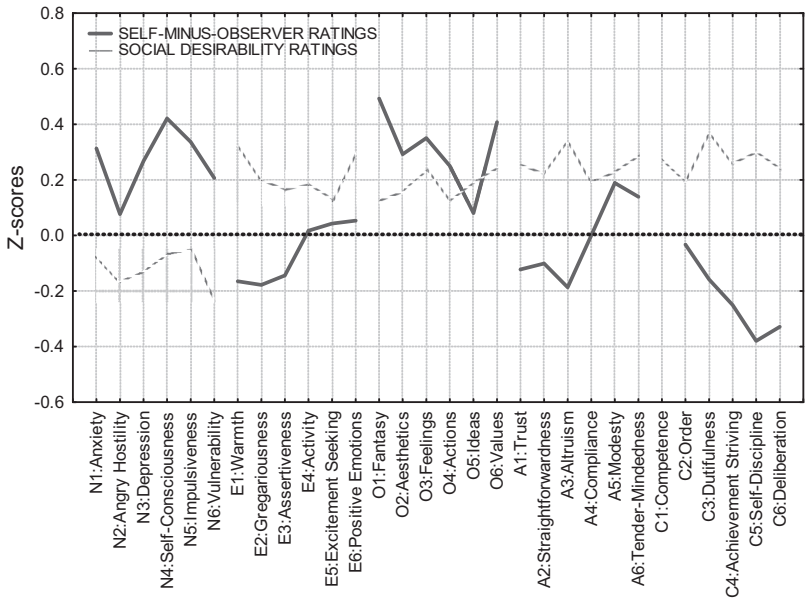

Figure 2. The average self-minus-observer profile for the four cultures studied in comparison with their average social desirability ratings on the Revised NEO Personality Inventory items. Letters–number combinations are the NEO PI-R facet scale numbers.

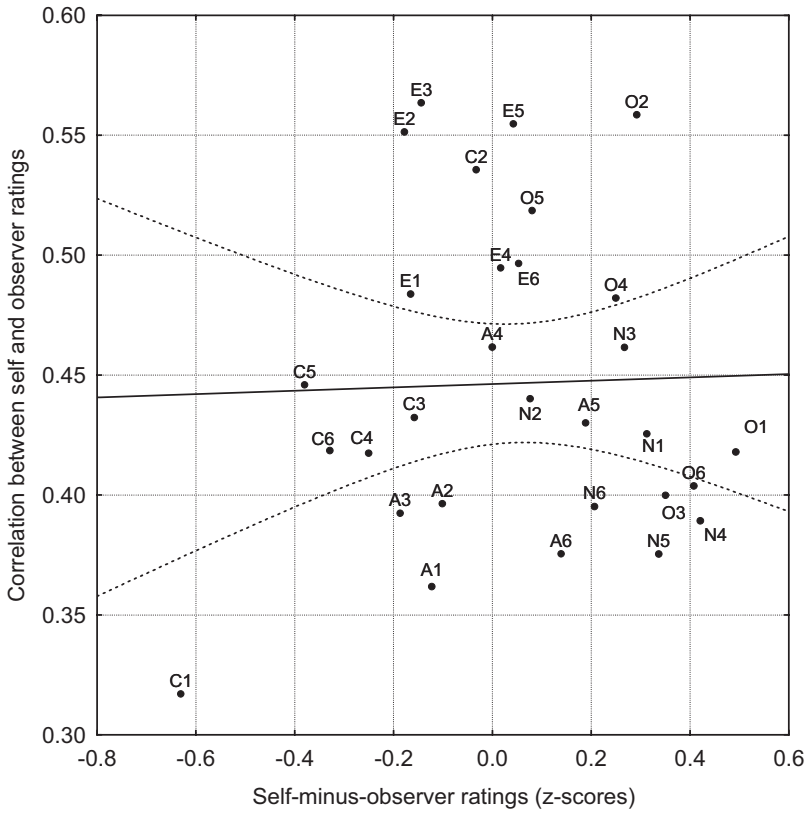

Figure 3. Correlation plot between self–other agreement and self-minus-observer-difference scores ($r = .03$, $p = .88$). Letters-number combinations are the NEO PI-R facet scale numbers.

(adults minus college age)—was positive ($r = .64$, $p < .001$), suggesting that the observer's perspective was more characteristic to a younger person, not an older person. Thus, an external observer may tend to emphasize the younglike personality trait levels of their targets, whereas generally younger targets may tend to show themselves as more maturelike.

Study 2

In spite of obvious language and cultural differences, all four studied countries are members of the European Union and have relatively high levels of human development and economic prosperity. Therefore, to make any claims concerning the universality of our findings in Study 1, we would need to have data from other regions of the World, including Asian and African countries. Disappointingly, the number of consensual studies between self- and observer-ratings done outside Europe and North America is in short supply. Fortunately, the NEO PI-R was translated into more than 40 languages and many researchers around the world have collected self-report data. In 2002, McCrae (2002) assembled self-report data that had been collected by other researchers using a variety of designs in 36 different cultures. Some cultures (e.g., Hong Kong) had only college-age respondents; some (e.g., Spain) had only adult data, and many had both. The ratio of men to women varied widely across cultures. In order to create overall culture scores that would be comparable across these diverse studies, McCrae (2002) first standardized each subsample using age- and gender-specific U.S. norms and then defined the overall culture *T* -score as the unweighted mean of all available subsamples. This strategy assumes that trait levels for a culture are generalizable across age and gender groups and that age and gender differences around the world are similar to those found in the U.S. norms. Both these assumptions were generally supported by the data (McCrae, 2002), and the validity of the resulting overall culture trait means was supported by their correlates (Allik & McCrae, 2004; McCrae & Terracciano, 2008).

Unlike self-report data, the collection of the observer's ratings has been much more systematic. During the Personality Profiles of Culture project, college students from 51 cultures identified an adult or college-aged man or woman whom they knew well and rated more than 12,000 targets using the R-form of the NEO PI-R (McCrae, Terracciano, & 78 Members of the Personality Profiles of Cultures Project, 2005; McCrae, Terracciano, & 79 Members of the Personality Profiles of Cultures Project, 2005). As there is a considerable overlap between the samples of nations for which aggregate self- or observer-rated personality scores are reported, it makes it possible to compare the self- and observer-rated personality traits at the aggregate national level. Provided that the pattern of difference between internal and external perspectives for the Big Five personality traits is pervasive, we could expect to observe it even if the targets of the self and observer's ratings are not identical, to say nothing about other differences (such as age, sex and occupation) between the study designs.

Method

The self-reported mean *T*-scores of the NEO PI-R subscales for 36 countries were published by McCrae (2002). An additional set of self-report data for Burkina Faso, Switzerland (French-speaking), and Poland was published by McCrae and Terracciano (2008, Appendix C). In another large-scale project, observers' ratings were collected from more than 50 different cultures using the R-Form of the NEO PI-R (McCrae, Terracciano, & 78 Members of the Personality Profiles of Cultures Project, 2005; McCrae, Terracciano, & 79 Members of the Personality Profiles of Cultures Project, 2005), meaning that standardized mean observer ratings for 51 cultures are now publicly available (McCrae & Terracciano, 2008). For 29 cultures, both self- and observer-reports are available. Although this overlapping set also contained data from Belgium, the Czech Republic, Estonia, and Germany, it is important to note that the data were different from what we used in Study 1. Because the published means in both cases are reported in *T*-scores, and the original data are not available, we converted them back to raw scores using the formula $([T\text{-Score} - 50] \cdot SD)/10 + M$, where *M* and *SD* are the mean and standard deviation of the U.S. adult or college age normative data (Costa & McCrae, 1992), dependent on the respective sample, and the average international sample data (McCrae & Terracciano, 2008) for self- and observer-ratings, respectively. It is important to note, however, that this back transformation from *T*-scores to raw scores is approximate, given that on the basis of *T*-scores alone it is impossible to reconstruct the exact scores for different sex and age groups. The reconstructed mean profile represents a hypothetical average person, without sex and age specification. To compute the self–other asymmetry index, we subtracted the mean score of the observer-ratings from the mean score of the self-ratings.

In order to study correlation with societal-level indicators, we found the mean absolute difference between observer-ratings and self-ratings for each culture. This score showing the magnitude of the self-minus-observer differences was correlated with several indicators characterizing economic and social conditions.

Gross domestic product (GDP). GDP at purchasing power parity in U.S. dollars, divided by the midyear population in 2006, were obtained from the Human Development Indices (2008).

Life expectancy. Life expectancy at birth indicates the number of years a newborn infant would live if prevailing patterns of age-specific mortality rates at the time of birth were to stay the same throughout the child's life (Human Development Indices, 2008).

Human Development Index (HDI). The Human Development Index measures the level of human development by combining normalized measures of life expectancy, literacy, educational attainment, and GDP per capita for countries worldwide; the reported indices are for the year 2006 (Human Development Indices, 2008).

Index of Shipping Difficulties. The Index of Shipping Difficulties is an indicator of required efforts and complications (border delays, fees, red tape, etc) met during shipping goods (World Development Report, 2009, Table A4).

Days to Start Business. The goal of the Doing Business project was to provide an objective basis for understanding and improving the regulatory environment for business. We used the days required for starting a business as an index of the bureaucratic and legal hurdles an entrepreneur must overcome to incorporate and register a new firm. Data were retrieved from http://doingbusiness.org/ExploreTopics/StartingBusiness/ on June 29, 2009.

Corruption Perception Index (CPI). The Corruption Perception Index ranks 180 countries by their perceived levels of freedom from corruption, as determined by expert assessments and opinion surveys. The CPI is compiled annually, and it was retrieved from the Transparency International homepage http://www.transparency.org/policy_research/surveys_indices/cpi on June 29, 2009.

Results and Discussion

The mean differential profiles (self-minus-observer) of the 30 NEO PI-R subscales for the 29 countries or cultural groups are shown in Table 1. The average between-country profile correlation was .43, suggesting that the profiles of the self-minus-observer mean differences are rather similar. The last column in Table 1 shows how much the mean self-minus-observer profile of each country is similar to the average self-minus-observer profile of all 29 countries. As nearly all correlations are positive and significant (median = .67), a strong first principal component is suggested on which all individual profiles are loading. There is only one country-level self-minus-observer profile of the 29 that clearly deviates from the common shape—the Danish profile. We have no good explanation for why the Danish data are conspicuously different from other countries. The Danish self- and observer-rating profiles alone are not outstanding from other profiles. The deviance of the self-minus-observer profile might reflect a real difference in perspective or might be a consequence of some measurement error, to say nothing about artifacts that could have been created by the back transformation from *T*-scores. Despite the ideosyncracy of the Danish self-minus-observer profile, we can still conclude that there is a remarkable cross-cultural similarity in the asymmetry of aggregate self- and observer-ratings.

The critical issue, however, is how well the culture-level findings of Study 2 agree with the individual-level findings of Study 1 that were obtained with a more controlled design. The cross-validity of the findings of the two studies is remarkable: Averaged self–observer difference profiles found in the four cultures investigated in Study 1 and the averaged self–observer difference profile of the 29 cultures investigated in Study 2 were correlated as highly as $r = .80$, $p < .0001$. Figure 4 presents the average self-minus-observer profile across 29 cultures, together with the average of four consensus studies (Study 1) and the differential profiles of the U.S. normative data (Costa & McCrae, 1992) for adults. All these three profiles are strongly correlated (from .70 to .82, $p < .0001$), suggesting that the average shape of the self-minus-observer profile remains essentially the same.

Further comparison of data from the Study 1 and Study 2 shows that the average self–observer differences profile of the 29 countries is not correlated to the visibility of traits, operationalized as the rank-order correlation between self- and observer-ratings in the pooled data of four European countries ($r = -.04$, $p = .83$). The lack of correlation is consistent with the findings in Study 1, further demonstrating that the self–observer asymmetry is probably not caused by the different amount of information available to self-raters and external observers. We also calculated the correla-

tion between the average self–observer differences profile of the 29 cultures and the mean social desirability ratings reported in Study 1. Similarly to the Study 1, the correlation was negative, but this time it was nonsignificant ($r = -.17$, $p = .37$). Thus, assuming that social desirability is relatively universal, the lack of significant positive correlation shows that in culture-level analyses, self–observer asymmetry cannot be explained by self-enhancement.

We were also interested whether the magnitude of the self-minus-observer differences is related to geographic, economic, and social indicators. In general, data of the 29 countries did not show significant correlation with these country-level indicators (Table 2), except a significant ($p = .008$) negative correlation with the days required for starting a business. Provided that it is not a statistical fluke, it remains to elucidate why, in countries with low bureaucratic and legal hurdles that an entrepreneur must overcome to incorporate and register a new firm, people generally see others more differently from how they see themselves.

General Discussion

Mainstream social psychology, focusing on human inabilities, has been engaged in expanding the list of errors in judgment (J. I. Krueger & Funder, 2004). Although this study is also about the disagreement between two perspectives from which personality can be judged, its message is, overwhelmingly, about the remarkable accuracy of personality judgments. First, the level of self–other agreement (the median rank-order correlation was .43) is noteworthy, considering the complicated chain of events required for an accurate personality judgment: The target of judgment must display behaviors and cues that are relevant to the trait being judged, and the judge must detect these cues and correctly process them (Funder & Colvin, 1997). Second, in addition to substantial agreement in ranking, the mean levels of personality trait judged by people and their acquaintances also converge (typically less than 0.2 units of a standard deviation), suggesting that the disparity between internal and external perspectives cannot be overly large, in contrast to the predictions of the actor–observer hypothesis (Malle, 2006; Malle, Knobe, & Nelson, 2007).

Third, in light of the smallness of the disparity between self- and observer-perspectives, it is noteworthy that the pattern of differences between self- and observer-ratings appears to be relatively systematic and universal, as it consistently transcends samples, languages, and measurement instruments. In four different European countries—Belgium, the Czech Republic, Estonia, and Germany—participants rated their 30 personality traits (i.e. the NEO PI-R model) using the same or very similar instrument and asked at least one other person to rate their personality traits using the same instrument. It appeared that although small in magnitude, the average differences between self-ratings and observer-ratings in the 30 personality traits were extremely similar in these four cultures. A secondary analysis of the culture-level mean scores of self- and observer-ratings obtained with the very same personality instrument in 29 cultures also revealed a replicable pattern self–observer differences. Most important, the culture-level profile of self–observer differences clearly replicated the individual-level differences found in the four European cultures; all of the self–observer differences also converged remarkably with differences that can be seen in the NEO PI-R normative data collected in the United States. All together, these findings provide reasonable grounds to suggest that albeit small, the self–observer asymmetry in personality descriptions is another nearly universal aspect of human personality description that remains relatively invariant, if not in all, then in the majority of human populations. It is very likely that there are variations, for instance related to sex, in the expression of this relatively universal pattern, but the evidence found so far speaks in favor of a universal trend. The existence of such universally replicable aspects of personality descriptions tells us that the descriptions truly reflect certain substantive features of human personality variation. Although there appears to be a small discrepancy between the self- and observer perspectives on personality traits, the systematic nature of the discrepancy may finally turn out to be extremely informative: Finding out why people see themselves as more neurotic and open, whereas others see them as more conscientious may finally lead to more refined insights into the processes of accurate person perception.

What are the main differences between internal and external perspectives? It seems that people almost everywhere see themselves as more neurotic and open, compared with how they are seen by other people. External observers generally hold a higher opinion of an individual's conscientiousness than he or she does about him- or herself. As a rule, people think that they are less assertive (E3) but much more excitement seeking (E5) and full of positive feelings (E6) than it seems from the vantage point of an external observer.

None of the existing theoretical schemes implying self–observer asymmetry has made specific predictions as to how people's descriptions of their own personality traits could be different from how they are described by their observers. In their daily lives, people regularly make personality judgments about themselves and other people. Although these judgments are not as systematic and comprehensive as personality questionnaires, they are supposedly of the same kind. Many social psychological theories endorsing actor–observer asymmetry seem to imply that in all personality judgments, voluntary or imposed, people are subjected to a pervasive disparity between self- and observer-descriptions. Undisputedly, individuals have the best perspective from which to judge personality characteristics relevant to internal states such as emotions (Beer & Watson, 2008; Spain et al., 2000). Because an external observer does not have privileged access to another's inner experience, the self–observer agreement on these traits could be compromised. It seems well established that a higher self–other agreement can be achieved on personality traits that are more visible or judge-able (Funder & Dobroth, 1987; Spain et al., 2000; Watson et al., 2000). Apparently, a logical consequence of this rule is that the self–observer disparity in mean ratings is also smaller on traits with higher visibility. Contrary to this prediction, however, in these studies there was no relationship between self-minus-observer scores and agreement between self- and observer-ratings on the 30 personality traits measured by the NEO PI-R. In other words, highly visible traits were not assessed more congruently in terms of mean levels by individuals and their judges.

Unexpectedly, both of the studies presented exclude any self-enhancement explanation (Kenny, 1994; Kwan et al., 2004) for the self–observer disparity in personality descriptions. In addition to the well-known fact that social desirability generally replicates the normative personality profile by amplifying its characteristic shape (Borkenau & Zaltauskas, 2009; Edwards, 1957), in this study we

Table 1
Self-Minus-Observer Ratings (Revised NEO Personality Inventory Facet Raw Scores) and Correlation With Other Profiles

Culture	N1	N2	N3	N4	N5	N6	E1	E2	E3	E4	E5	E6	O1	O2	O3	O4	O5
Austria	2.2	0.3	1.1	2.1	1.4	0.6	0.1	0.1	−1.8	−0.4	0.9	1.6	6.1	5.1	4.2	3.5	2.5
Belgium	1.1	−0.7	0.8	2.7	0.0	0.0	−1.2	−0.3	−2.3	−0.2	0.8	0.3	1.8	2.9	1.8	1.5	1.1
Burkina Faso	3.0	0.4	3.0	3.0	0.8	1.7	−0.5	−0.5	−1.7	−0.8	1.1	−0.5	1.2	2.8	0.2	2.0	2.2
Canada	−2.6	−3.2	−1.9	0.5	0.1	−3.1	0.2	−3.5	−2.2	1.3	−1.2	1.5	0.9	2.9	1.5	2.8	1.4
China	1.4	1.9	1.7	2.9	0.9	1.1	1.0	1.0	−2.4	0.2	1.7	0.1	0.8	2.8	1.5	−0.9	0.5
Croatia	0.4	−0.6	1.0	2.3	−1.2	−0.5	−0.1	−1.1	−3.0	0.5	0.4	−0.5	1.0	3.7	1.0	0.3	1.7
Czech Republic	0.2	−1.3	1.0	0.7	2.2	2.1	−0.5	0.2	−0.7	0.4	2.2	2.6	4.7	1.6	2.1	3.1	1.6
Denmark	−3.6	−3.7	−4.1	−1.2	−1.3	−3.9	0.1	−0.2	−1.9	0.4	−0.5	1.1	−2.8	−2.3	−1.3	0.9	−1.7
Estonia	−0.3	−2.0	−0.4	0.5	1.0	−1.7	1.2	−1.7	−1.0	−0.3	2.0	1.3	4.4	4.8	4.0	2.9	3.1
France	0.2	−1.5	0.8	1.0	1.2	−0.5	1.5	0.8	−0.9	−0.1	1.3	1.9	2.5	3.0	2.8	3.0	2.4
Germany	1.9	0.7	2.0	2.3	2.1	1.6	−1.0	−0.3	−1.8	−0.3	0.2	1.2	3.5	1.8	2.0	2.5	0.6
Hong Kong	2.5	−1.8	1.8	4.2	1.2	2.5	0.2	−1.1	−2.5	0.2	0.5	0.1	2.5	4.1	1.5	1.1	1.1
India	1.7	2.9	1.6	1.3	−0.2	1.9	1.0	−0.7	−3.9	0.5	3.8	−0.8	−2.1	2.1	0.8	0.7	0.9
Indonesia	0.6	−1.5	0.1	1.0	1.8	−0.7	2.0	1.6	−0.8	1.0	2.8	2.5	3.4	3.6	2.3	1.2	1.8
Italy	0.7	−0.9	0.3	0.4	0.5	−0.9	2.3	1.6	−0.5	1.3	1.3	1.5	2.3	3.3	2.8	0.6	1.0
Japan	2.3	0.0	2.4	2.7	1.8	2.3	−1.9	−1.9	−3.5	−2.6	−0.4	−0.8	2.8	1.0	0.8	1.1	1.4
Korea, Republic	3.3	1.7	3.7	4.3	1.4	2.9	−2.3	−1.2	−3.5	−0.5	1.0	−2.2	2.9	3.0	0.5	−0.4	1.1
Malaysia	2.2	−0.7	2.4	2.5	1.1	0.9	−0.1	−1.6	−1.9	−0.6	0.5	0.8	1.4	2.5	1.6	2.6	1.3
Peru	3.2	0.9	2.3	1.3	1.3	1.0	0.4	0.6	−0.6	0.0	2.5	0.9	3.5	3.7	2.4	0.2	1.3
Philippines	2.2	0.5	2.8	2.6	2.8	2.2	0.5	−0.2	−2.1	−0.3	1.8	1.7	2.0	3.6	1.2	0.0	1.7
Poland	1.0	−0.5	1.4	1.7	−1.1	0.6	−0.3	−0.4	−2.4	1.2	0.6	0.5	0.8	2.3	0.8	1.5	0.4
Portugal	1.4	−0.3	1.7	2.2	0.0	0.4	−0.4	−1.3	−2.4	−0.5	0.8	−0.4	0.0	2.9	0.4	1.7	0.1
Russia	0.5	−0.1	1.7	0.6	0.5	0.1	0.8	1.4	−0.6	0.9	1.0	2.2	2.2	2.3	1.8	2.3	0.7
Serbia	0.0	−1.8	−0.1	0.1	0.2	−0.6	1.1	0.7	−1.0	0.9	0.9	1.4	2.9	5.1	2.9	2.7	3.1
Spain	3.2	1.0	3.1	2.5	1.5	2.3	−1.5	−0.6	−3.5	0.4	3.1	−0.2	2.9	2.0	1.7	0.6	0.9
Switzerland (German)	−1.5	−1.3	−2.2	1.6	0.6	−0.3	−0.9	0.9	−0.2	0.4	1.8	2.0	3.4	4.8	4.2	3.3	3.9
Switzerland (French)	4.4	2.2	4.5	2.7	3.5	2.4	0.1	0.0	−2.1	−0.4	3.5	0.5	3.3	0.2	0.3	0.8	−1.1
Turkey	1.2	0.5	1.5	1.7	1.6	0.3	−0.1	1.7	−1.5	0.8	4.2	1.9	3.5	3.5	3.0	2.5	2.7
United States	0.0	−1.2	0.1	1.5	1.0	−0.7	0.8	−1.2	−1.5	0.9	0.0	0.7	1.0	1.7	1.6	2.0	0.7
Mean	1.1	−0.3	1.2	1.8	0.9	0.5	0.1	−0.2	−1.9	0.1	1.3	0.8	2.2	2.8	1.7	1.6	1.3

Note. r = Correlation with the mean self-minus-observer profile of 29 cultures. Significant correlations ($p < .05$) are $r > |.36|$. Letter-number combinations are the NEO PI-R facet scale numbers.

provided solid empirical evidence that social desirability ratings strongly converge across different cultures. Thus, similarly to the normative personality profiles, social desirability ratings seem to share a core common across languages and cultures. Powerful evidence of such pancultural agreement has been previously provided by Williams, Satterwhite, and Saiz (1998), who gathered favorability ratings of trait adjectives in 10 diverse nations and found between-country correlations ranging from .68 to .95 (*Mdn* = .82). For example, in all cultures examined, *friendly* and *responsible* were highly favorable, *ambitious* was somewhat favorable, and *arrogant* was unfavorable. Based on this, it was important that the average self-minus-observer profile was negatively correlated with social desirability ratings in both individual-level and culture-level data. Instead of self-enhancement, there is more reason to talk about self-deprecation: People seem to view themselves less favorably than they view others or are perceived by others, at least outside tightly controlled social psychological experiments.

However, the finding that it is the observer, rather than the self, who enhances personality descriptions is less surprising considering the typical recruitment procedure. Since observers need to know their targets well, those who consent to participate are not only familiar with their targets but are in many cases intimately related to or involved with them. Very often they are close relatives (parents or children), spouses, or very good friends who have known each other for many years. It is interesting to note that the actor–observer asymmetry holds only when the actor and the observer are intimates, whereas no asymmetry emerges when the actor and the observer are strangers (Malle, 2006). This principal study design, where the actor and the observer are intimates, almost inevitably leads to observers initially having more positive attitudes towards their targets than they might have towards a complete stranger. In other words, there is no doubt that, in most cases, targets belonged to the observer's in-groups, and his or her judgment might have been affected, to a lesser or larger extent, by intimate relationship or in-group favoritism (Hewstone, Rubin, & Willis, 2002; Mullen, Brown, & Smith, 1992).

Although the negative correlation with social desirability easily leads to the hypothesis that all judges rated their friends and therefore the observed disparity is not about self–other perspective difference but about self-friend perspective difference, the "friend-enhancement" (or "self-deprecation") may be not the best, or at least only, explanation for the systematic self–observer disparity in the mean scores. Namely, self-rating scores are higher on neuroticism, which is indeed consistent to the "friend-enhancement" explanation, but they are also higher on Openness, which is contrary to the friend-enhancement hypothesis. There are also within-domain facet differences (like the big difference on C1: Competence vs. C2: Order or E3: Assertiveness vs. E5: Excitement Seeking or A3: Altruism vs. A5: Modesty) that cannot

O6	A1	A2	A3	A4	A5	A6	C1	C2	C3	C4	C5	C6	*r*
3.5	−0.4	−1.7	−0.5	−0.5	0.7	1.7	−2.1	−2.2	−1.8	−0.6	−4.1	−2.0	.86
2.5	0.7	0.1	0.2	1.4	1.9	2.0	−0.1	0.4	1.2	−0.8	−0.6	−0.2	.88
0.9	−1.1	−0.1	0.5	0.9	1.7	1.7	−1.4	0.0	−0.1	1.2	−1.9	−0.3	.77
3.3	4.3	3.6	2.2	3.2	1.6	2.5	1.5	0.6	2.0	−0.4	1.1	−0.4	.35
1.9	2.5	1.0	1.3	−2.5	2.2	2.7	−0.5	0.0	2.2	0.8	0.1	1.3	.51
0.2	0.1	−0.3	0.7	1.3	0.6	1.8	−0.2	0.3	1.8	0.5	0.1	1.0	.66
1.7	−1.0	−1.0	−0.9	0.1	−0.2	−0.1	−2.2	0.4	−1.0	−0.8	−1.8	−1.1	.69
−0.6	2.7	2.1	0.6	3.0	2.2	−0.1	−0.5	1.9	2.0	0.4	2.6	0.3	−.22
1.1	2.7	0.2	0.8	1.5	4.6	3.7	−0.9	−0.1	2.1	−0.2	0.1	0.2	.73
1.7	0.1	−0.4	1.1	0.9	2.6	1.9	−1.0	0.6	1.2	0.4	−1.0	0.0	.82
2.2	−0.9	−1.2	−0.8	−0.3	0.1	1.4	−2.7	−1.0	−2.2	−1.3	−3.2	−2.3	.80
2.8	1.6	4.3	1.0	4.2	2.6	2.0	−1.6	−1.4	−0.4	−0.5	−1.0	0.2	.76
−0.4	0.4	1.7	2.0	−0.7	2.6	2.4	−1.0	−0.9	0.2	−0.7	−1.4	0.3	.49
1.5	2.1	1.3	1.5	1.1	1.5	0.6	−0.7	0.2	0.5	1.4	−1.7	0.4	.71
1.0	1.4	2.8	2.2	1.0	2.2	0.8	0.2	−0.3	2.4	0.4	0.4	1.2	.47
1.6	1.0	0.0	−0.2	0.7	1.8	0.6	−3.7	−0.8	−1.2	−2.5	−3.6	−0.4	.81
2.5	−0.1	0.3	0.7	−0.6	1.5	1.6	−2.0	−0.8	1.5	0.5	−3.8	−0.2	.73
−1.3	1.4	−0.1	2.1	1.8	1.6	4.0	−1.5	1.3	0.2	0.3	−2.9	−0.6	.78
1.7	−0.2	−0.8	0.1	−0.3	1.3	1.6	−0.2	−0.9	0.5	0.6	−3.1	−0.2	.78
3.4	3.0	1.5	0.5	2.5	2.6	2.6	−2.1	−0.5	−1.0	−1.5	−2.4	0.0	.82
1.3	0.7	0.3	0.7	1.4	2.7	1.2	−0.8	0.5	1.2	0.0	−0.1	−0.7	.77
−0.8	−0.2	0.0	0.0	1.8	1.6	1.3	−1.6	0.1	0.5	−0.3	−1.2	0.0	.78
0.2	0.0	−1.4	−0.4	0.7	1.1	−0.2	−1.2	0.1	−0.3	−0.9	−0.8	−1.5	.68
1.2	0.9	1.1	1.6	1.2	1.8	0.9	−1.5	0.6	0.5	0.4	−1.3	−0.4	.76
0.7	0.2	−1.1	0.3	−1.0	0.5	3.2	−2.3	−1.0	−0.8	−0.4	−3.2	−0.6	.80
2.9	0.9	−3.6	−0.9	−1.6	−1.7	0.1	1.3	1.2	0.1	0.7	−1.5	0.2	.48
0.5	−1.3	−0.2	−1.2	−0.2	2.8	2.4	−4.7	−2.0	−2.2	−1.6	−4.0	−2.6	.64
2.2	−0.9	−1.3	−0.1	−0.4	−1.4	−0.6	−0.2	−2.3	−0.9	−0.2	−1.7	−0.6	.66
1.6	1.8	1.1	1.2	1.6	1.2	1.8	0.5	0.8	1.8	−0.1	0.2	−0.1	.62
1.4	0.8	0.3	0.6	0.8	1.5	1.6	−1.1	−0.2	0.4	−0.2	−1.4	−0.3	.67

be explained by the "friend-enhancement" hypothesis (nor by the self-enhancement hypothesis).

Obviously, the broad explanations based on visibility and social desirability of personality traits are not the only possible accounts for the asymmetries between the self and observer perspective. Cognitive psychologists have documented several subtle differences in the ways people process information about themselves and about others. For example, memories recalled from the first-person position are phenomenologically more rich and emotionally more intensive, whereas memories from the third-person position are more intellectual and schematic (Nigro & Neisser, 1983). Although the knowledge about representational models on which personality judgments are based is still limited, it is quite clear that the information about one's own personality traits is largely represented at the level of abstract semantic knowledge, rather than in the form of recollections of individual behaviors exemplifying these personality traits (Klein & Loftus, 1993). It is remarkable that a consensually accurate knowledge of personality traits can be achieved even in the complete absence of episodic memory and ability to recollect any specific behavioral instances (Tulving, 1993). Nevertheless, we judge others more on what we see (how the subjects looked or what they did) but ourselves more often than not based on what we think, feel, or intend to do (McIsaac & Eich, 2002; Pronin, 2008). Evaluation of different personality traits may also encourage participants to retrieve information selectively, so that content varies on the dimensions such as thinking versus experiencing or semantic versus episodic. For instance, Neuroticism can be best estimated on the basis of recollection of affective episodes, Conscientiousness can be best estimated on the basis of recording behavioral incidences, and Openness can be best estimated mainly on the basis of reasoning about the target's interests, ideas, and values. Indeed, when a group of raters were asked to rate the content of the NEO PI-R items, those that were measuring Conscientiousness were rated highest in terms of directly observable actions, Openness items were on the top of thinkinglike ratings, and Neuroticism items were maximally saturated by the emotional content (Pytlik Zillig, Hemenover, & Dienstbier, 2002). Thus, Neuroticism and Openness can be considered largely as intrapsychic dimensions, defined in terms of characteristics of consciousness (McCrae, 1996). It seems that, in all places, people tend to think that they are more curious and that their lives are experientially richer than is perceived from an external observer's perspective. They also perceive their neurotic tendencies to be more salient than they are perceived from the observer's perspective. On the other hand, Assertiveness and Conscientiousness can be assessed on the basis of overt actions, which give external observers more instances of these traits than to targets who focus on their thoughts and feelings. Consequently, persons perceive others as more competent and self-disciplined than the targets perceive themselves. However, it is not clear how exactly the

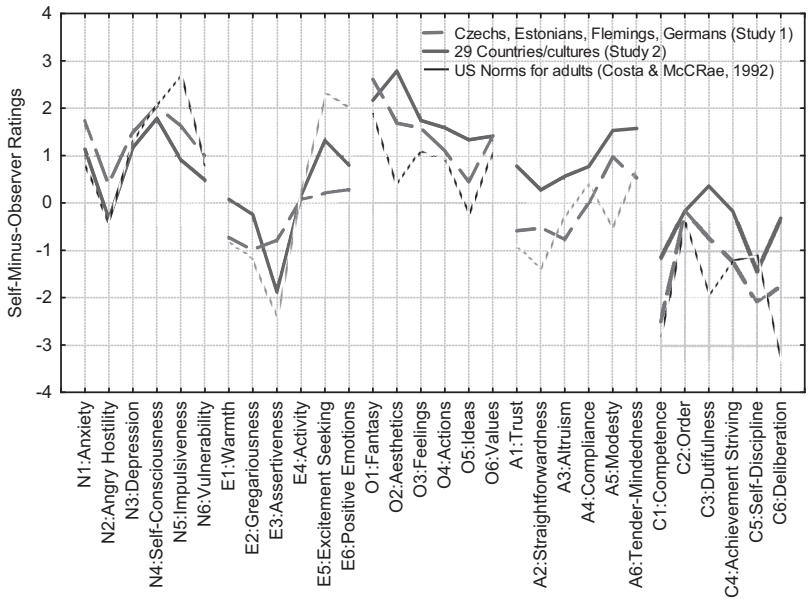

Figure 4. The average self-minus-observer profile (raw scores) for four consensus studies (Study 1), 29 world cultures (Study 2), and U.S. normative data (Costa & McCrae, 1992) for adults. Letters–number combinations are the NEO PI-R facet scale numbers.

psychological content of personality traits differs from their mere visibility, which is supposed to regulate the degree of their self–observer agreement. In the absence of decisive data we can only speculate about cognitive mechanisms that could be responsible for the delicate differences in how people see others and how they see themselves.

To conclude, it is certainly significant that the result of this cross-cultural inquiry failed to confirm any social psychological theories about self-enhancement or the actor–observer hypothesis. Although the results presented cannot refute the existence of the actor–observer hypothesis in the way that a careful meta-analysis certainly can (Malle, 2006), it is still possible to demonstrate that when a large number of persons are asked to describe their own or somebody else's personality, there is no evidence of a significant actor–observer asymmetry or of self-enhancement. In addition to pointing to the need to test social psychological theories more rigorously outside the laboratory in more realistic settings, this study has also revealed a need for future research to provide a satisfactory explanation for a small but a cross-culturally replicable disparity between self- and observer-perspectives.

Table 2
Correlation Between the Size of the Self-Minus-Observer Differences and Societal-Level Indicators

Indicator	*r*	*N*	*p*
Latitude	.02	29	.911
Longitude	−.26	29	.171
Human Development Index (2006)	.16	27	.417
Life expectancy at birth (2006)	.27	27	.174
GDP at PPP (2006)	.24	27	.235
Democracy (rank)	−.18	27	.366
Index of Shipping Difficulties	**−.50**	27	.008
Day to start business	−.04	27	.844
Corruption Perception Index	.42	27	.028

Note. The significant correlation $p < .01$ is shown in bold. GDP = gross domestic product; PPP = purchasing power parity in U.S. dollars.

References

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49,* 1621–1630. doi:10.1037/0022-3514.49.6.1621

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology, 68,* 804–825. doi:10.1037/0022-3514.68.5.804

Allik, J., & McCrae, R. R. (2004). Toward a geography of personality traits—Patterns of profiles across 36 cultures. *Journal of Cross-Cultural Psychology, 35,* 13–28. doi:10.1177/0022022103260382

Allik, J., Realo, A., Mõttus, R., Pullmann, H., Trifonova, A., & McCrae, R. R. (2009). Personality traits of Russians from the observer's perspective. *European Journal of Personality, 23,* 567–588. doi:10.1002/per.721

Beer, A., & Watson, D. (2008). Asymmetry in judgments of personality: Others are less differentiated than the self. *Journal of Personality, 76,* 535–560. doi:10.1111/j.1467-6494.2008.00495.x

Borkenau, P., & Zaltauskas, K. (2009). Effects of self-enhancement on agreement on personality profiles. *European Journal of Personality, 23,* 107–123. doi:10.1002/per.707

Colvin, C. R. (1993). "Judgable" people: Personality, behavior, and competing explanations. *Journal of Personality and Social Psychology, 64,* 861–873. doi:10.1037/0022-3514.64.5.861

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,* 110–117. doi:10.1111/j.1468-2389.2007.00371.x

Costa, P. T., & 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., Martin, T. A., Oryol, V. E., Senin, I. G., Rukavishnikov, A. A., . . . Realo, A. (2000). Personality development from adoloscence through adulthood: Further cross-cultural comparisons of age differences. In V. J. Molfese & D. L. Molfese (Eds.), *Temperament and personality development across life span* (pp. 235–252). Mahwah, NJ: Erlbaum.

Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81,* 322–331. doi:10.1037/0022-3514.81.2.322

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* New York, NY: Dryden.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116,* 429–456. doi:10.1037/0033-2909.116.3.429

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

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55,* 149–158. doi:10.1037/0022-3514.55.1.149

Funder, D. C., & Colvin, C. R. (1997). Congruence of others' and self-judgments on personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 617–647). San Diego, CA: Academic Press.

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52,* 409–418. doi:10.1037/0022-3514.52.2.409

Hřebíčková, M. (2002). Internal consistency of the Czech version of the NEO Personality Inventory (NEO-PI-R). *Ceskoslovenska Psychologie, 46,* 521–535.

Heine, S. J., & Hamamura, T. (2007). In search of East Asian self-enhancement. *Personality and Social Psychology Review, 11,* 4–24. doi:10.1177/1088868306294587

Heine, S. J., & Renshaw, K. (2002). Interjudge agreement, self-enhancement, and liking: Cross-cultural divergences. *Personality and Social Psychology Bulletin, 28,* 578–587. doi:10.1177/0146167202288002

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53,* 575–604. doi:10.1146/annurev-.psych.53.100901.135109

Hoekstra, H. A., Ormel, J., & DeFruyt, F. (1996). *NEO Persoonlijkheidsvragenlisten NEO-PI-R en NEO-FF-I. Handleigin* [*Revised NEO Personality Inventory (NEO PI-R) and NEO-FFI manual*]. Lisse, the Netherlands: Swets & Zeitlinger.

Human Development Indices. (2008). *A statistical update 2008.* New York, NY: The United Nations Development Program.

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,* 521–551. doi:10.1111/j.1467-6494.1993.tb00781.x

Jones, E. E., & Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior.* Morristown, NJ: General Learning Press.

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. doi:10.1002/1099-0984(200005/06)14:3<265::AID-PER376>3.0.CO;2-B

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York, NY: Guilford Press.

Klein, S. B., & Loftus, J. (1993). The mental representation of trait and autobiographical knowledge about the self. In T. K. Srull & R. S. Wyer (Eds.), *The mental representation of trait and autobiographical knowledge about the self* (pp. 1–50). Hillsdale, NJ: Erlbaum.

Konstabel, K., Aavik, T., & Allik, J. (2006). Social desirability and consensual validity of personality traits. *European Journal of Personality, 20,* 549–566. doi:10.1002/per.593

Krueger, J. (1998). Enhancement bias in descriptions of self and others. *Personality and Social Psychology Bulletin, 24,* 505–516. doi:10.1177/0146167298245006

Krueger, J. I., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences, 27,* 313–327. doi:10.1017/S0140525X04000081

Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., & Robins, R. W. (2004). Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. *Psychological Review, 111,* 94–110. doi:10.1037/0033-295X.111.1.94

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,* 1001–1053.

Locke, K. D. (2002). Are descriptions of the self more complex than descriptions of others? *Personality and Social Psychology Bulletin, 28,* 1094–1105. doi:10.1177/01461672022811008

Malle, B. F. (2006). The actor–observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological Bulletin, 132,* 895–919. doi:10.1037/0033-2909.132.6.895

Malle, B. F., Knobe, J. A., & Nelson, S. E. (2007). Actor–observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology, 93,* 491–514. doi:10.1037/0022-3514.93.4.491

McCrae, R. R. (1996). Social consequences of experiential openness. *Psychological Bulletin, 120,* 323–337. doi:10.1037/0033-2909.120.3.323

McCrae, R. R. (2002). NEO-PI-R data from 36 cultures: Further intercultural comparisons. In R. R. McCrae & J. Allik (Eds.), *The Five-Factor Model of personality across cultures* (pp. 105–125). New York, NY: Kluwer Academic/Plenum.

McCrae, R., & Costa, P. (1997). Personality trait structures as a human universal. *American Psychologist, 52,* 509–516.

McCrae, R. R., Costa, P. T., Hřebíčková, M., Urbanek, T., Martin, T. A., Oryol, V. E., . . . Senin, I. G. (2004). Age differences in personality traits across cultures: Self-report and observer perspectives. *European Journal of Personality, 18,* 143–157. doi:10.1002/per.510

McCrae, R. R., Costa, P. T., Jr., Pedroso de Lima, M., Simoes, A., Ostendorf, F., Angleitner, A., . . . Piedmont, R. L. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology, 35,* 466–477. doi:10.1037/0012-1649.35.2.466

McCrae, R. R., Costa, P. T., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., . . . Urbánek, T. (2004). Consensual Validation of Personality Traits Across Cultures. *Journal of Research in Personality, 38,* 179–201. doi:10.1016/S0092-6566(03)00056-4

McCrae, R. R., & Terracciano, A. (2008). The Five-Factor Model and its correlates in individuals and cultures. In F. van de Vijver, D. A. van Hemert, & Y. H. Poortinga (Eds.), *Multilevel analysis of individuals and cultures* (pp. 249–283). New York, NY: Erlbaum.

McCrae, R. R., Terracciano, A., & 78 Members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88,* 547–561. doi:10.1037/0022-3514.88.3.547

McCrae, R. R., Terracciano, A., & 79 Members of the Personality Profiles of Cultures Project. (2005). Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology, 89,* 407–425. doi:10.1037/0022-3514.89.3.407

McIsaac, H. K., & Eich, E. (2002). Vantage point in episodic memory. *Psychonomic Bulletin & Review, 9,* 146–150.

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,* 155–170. doi:10.1016/j.jrp.2006.03.007

Mõttus, R., Pullmann, H., & Allik, J. (2006). Toward more readable Big Five personality inventories. *European Journal of Psychological Assessment, 22,* 149–157. doi:10.1027/1015-5759.22.3.149

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22,* 103–122. doi:10.1002/ejsp.2420220202

Nigro, G., & Neisser, U. (1983). Point of view in personal memories. *Cognitive Psychology, 15,* 467–482. doi:10.1016/0010-0285(83)90016-6

Nisbett, R. E., Caputo, C., Legant, P., & Marecek, J. (1973). Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology, 27,* 154–164. doi:10.1037/h0034779

Ostendorf, F., & Angleitner, A. (2004). *NEO-PI-R: NEO-Persönlichkeitsinventar nach Costa und McCrae* [NEO Personality Inventory according to Costa and McCrae]. Göttingen, Germany: Hogrefe.

Pronin, E. (2008, May). How we see ourselves and how we see others. *Science, 320,* 1177–1180. doi:10.1126/science.1154199

Pytlik Zillig, L. M., Hemenover, S. H., & Dienstbier, R. A. (2002). What do we assess when we assess a Big 5 trait? A content analysis of the affective, behavioral, and cognitive processes represented in Big 5 personality inventories. *Personality and Social Psychology Bulletin, 28,* 847–858. doi:10.1177/0146167202289013

Sande, G. N., Goethals, G. R., & Radloff, C. E. (1988). Perceiving ones own traits and others: The multifaceted self. *Journal of Personality and Social Psychology, 54,* 13–20. doi:10.1037/0022-3514.54.1.13

Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martinez, V. (2007). The geographic distribution of big five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology, 38,* 173–212. doi:10.1177/0022022106297299

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology, 89,* 623–642. doi:10.1037/0022-3514.89.4.623

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in big five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94,* 168–182. doi:10.1037/0022-3514.94.1.168

Spain, J. S., Eaton, L. G., & Funder, D. C. (2000). Perspectives on personality: The relative accuracy of self versus others for the prediction of emotion and behavior. *Journal of Personality, 68,* 837–867. doi:10.1111/1467-6494.00118

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology, 84,* 1041–1053. doi:10.1037/0022-3514.84.5.1041

Tulving, E. (1993). Self-knowledge of an amnesiac patient is represented abstractly. In T. K. Srull & R. S. Wyer (Eds.), *The mental representation of trait and autobiographical knowledge about the self* (pp. 147–156). Hillsdale, NJ: Erlbaum.

Van Hemert, D. A., van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2002). Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries. *Personality and Individual Differences, 33,* 1229–1249.

Watson, D. (1982). The actor and the observer: How are their perceptions of causality divergent. *Psychological Bulletin, 92,* 682–700. doi:10.1037/0033-2909.92.3.682

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,* 546–558. doi:10.1037/0022-3514.78.3.546

Williams, J. E., Satterwhite, R. C., & Saiz, J. L. (1998). *The importance of psychological traits: A cross-cultural study.* New York, NY: Plenum Press.

World Development Report. (2009). *Reshaping economic geography.* Washington, DC: The World Bank.

Received June 3, 2009
Revision received February 8, 2010
Accepted February 11, 2010 ■