Published in final edited form as:
J Res Pers. 2013 April ; 47(2): 135–144. doi:10.1016/j.jrp.2012.12.001.

Do Men Vary More than Women in Personality? A Study in 51 Cultures

Peter Borkenau,
Department of Psychology, Martin-Luther Universität Halle-Wittenberg, Halle, Germany

Robert R. McCrae, and
Baltimore, MD, USA

Antonio Terracciano
Department of Geriatrics, College of Medicine, Florida State University, Tallahassee, FL, USA

National Institute of Aging, Baltimore MD, USA

Abstract

Do men vary more than women in personality? Evolutionary, genetic, and cultural arguments suggest that hypothesis. In this study we tested it using 12,156 college student raters from 51 cultures who described a person they knew well on the 3rd-person version of the Revised NEO Personality Inventory. In most cultures, male targets varied more than female targets, and ratings by female informants varied more than ratings by male informants, which may explain why higher variances for men are not found in self-reports. Variances were higher in more developed, and effects of target sex were stronger in more individualistic societies. It seems that individualistic cultures enable a less restricted expression of personality, resulting in larger variances and particularly so among men.

Keywords

cross-cultural research; individualism; multilevel modeling; sex differences; variability

Numerous studies have shown that men vary more than women in cognitive abilities (Arden & Plomin, 2006; Benbow, 1988; Deary, Thorpe, Wilson, Starr, & Whally, 2003; Feingold, 1995; Hedges & Friedman, 1993; Hedges & Nowell, 1995; Humphreys, 1988). Moreover, higher intrasex variances in men have been found for heterogeneous sets of variables like birth weight, adult height, 60-meter dash times, and numerous blood parameters (Lehre, Lehre, Laake, & Danbolt, 2009). Thus more variability between men than women seems to be a quite widespread phenomenon, raising the question whether this applies to personality too.

There are several reasons to expect that men vary more than women in personality. Evolutionary psychologists have argued that a larger variety of qualities is compatible with

Correspondence concerning this article should be addressed to Peter Borkenau, Department of Psychology, Martin-Luther University Halle-Wittenberg, D-06099 Halle, Germany. p.borkenau@psych.uni-halle.de.

reproductive success among men, as females but not males are restricted to high parental investment strategies. This may have given rise to greater male variability in sexually selected traits (Archer & Mehdikhani, 2003).

Furthermore, there are genetic hypotheses: Candidate genes to account for larger male variability may be located on the X chromosome, on which many genes important for the nervous system have been identified (Check, 2005). As females have two X chromosomes whereas males have only one, and genetic effects on personality seem to be additive to some extent (Krueger & Johnson, 2008), the effects of two X-linked alleles may be averaged in females, resulting in less extreme phenotypes unless the female is homozygous. That holds although most regions on one of the two X chromosomes in females are inactivated, because: (a) some regions of the X chromosome escape that inactivation; and (b) for the remaining regions a mosaicism is typical in human females, that is, the maternal X-linked allele remains activated in about one half of the cells and the paternal allele in the other half (Amos-Landgraf et al., 2007). This implies that some kind of averaging of maternal and paternal X-linked genetic effects may occur across cells. By contrast, males inherit one X-chromosome only, and its effects are fully expressed resulting in more extreme phenotypes and consequently higher variances.

Social-role explanations for a higher variability in men are reasonable as well: Various social norms probably are – or at least have been – stricter for women, allowing men to choose between more diverse behavioral options. But evidence to support that argument is currently sparse as research on gender roles has focused on sex differences in mean levels – e.g., that men are supposed to be more agentic whereas women are supposed to be more communal (Bosak, Sczesny, & Eagly, 2012) – not on sex differences in the diversity of behavioral options.

That men vary more than women in personality is not only reasonable; investigating that is also important. Liability-threshold or continuity models of mental disorders imply that extreme trait levels tend to be abnormal. Therefore, a higher variance in a group should result in a higher proportion of its members exceeding a “critical” threshold and being classified as abnormal (Hedges & Friedman, 1993). Thus if there were no sex differences in mean levels but men would vary more than women, the proportion of group members exceeding that threshold would be higher in men. For example, that men are strongly overrepresented among prison inmates may not mainly reflect a large sex difference in average levels of antisocial behavior, but that more men than women manifest extreme levels of antisocial behavior, relative to their intragroup means. Moreover, effects of sex differences in mean levels and in variances on the proportion of individuals exceeding a threshold may add up (Feingold, 1995; Humphreys, 1988). Thus a moderately higher mean level (Eagly & Steffen, 1986) in combination with a possibly higher variance in men’s aggressiveness may account for their disproportionate crime rate compared to that of women.

So what is the evidence on sex differences in variability in personality? There are only a few observational studies on this issue, among them a study by Mehl, Vazire, Ramirez-Esparza, Slatcher, & Pennebaker (2007). In six samples, these authors analyzed sex differences in talkativeness measured by an electronically activated recorder. They did not find systematic sex differences in average talkativeness, but the standard deviations of estimated words spoken per day were higher for men than for women in each of the six samples. If there were no sex differences in variability in the population, the probability of such a finding would be less than 2%, according to a binomial test. Thus this observational study suggests that measured talkativeness varies more among men than among women.

More evidence is available on self-reported sex differences in personality. First in the manuals of many self-report instruments, means and standard deviations are reported separately for women and for men. Second, there are several meta-analyses on sex-differences in personality, although most of them analyzed only the means and used the variances for obtaining effect size estimates only. But that is changing. For example, Cross, Copping, and Campbell (2011) published a meta-analysis on sex differences in impulsivity in adults, analyzing mean levels as well as variances. These authors expected more variability between men than between women, but did not find it except for the disinhibition facet of the Sensation Seeking Scale. The authors explain this nonconfirmation of their expectations by a sampling bias: Clinical and incarcerated samples were excluded, and given the overrepresentation of men in pathological and criminal behavior in which risk taking is a factor, this constraint may have reduced the male more than the female variance.

Another explanation, however, that Cross et al. (2011) do not discuss is that in most of the studies in their meta-analysis personality was measured by self-report, where the sex of the person being described (the target) and the sex of the person who provides the description (the informant) are entirely confounded. Effects of target sex might therefore be masked by countervailing effects of informant sex. But effects of target sex and of informant sex can be separated in descriptions by knowledgeable informants. Borkenau, Hřebíčková, Kuppens, Realo, and Allik (2013) therefore compared the intrasex variances in self-reports and in informant reports of personality, measured by either the NEO PI-R (Costa & McCrae, 1992) or the NEO PI-3 (McCrae & Costa, 2010), in four samples from Belgium, the Czech Republic, Estonia, and Germany. Whereas self-reports did not vary more among men than among women, the variances were systematically higher for male targets in the descriptions by informants, Neuroticism constituting an exception. In addition, a countervailing effect was obtained for informant sex, in that descriptions by women varied more than descriptions by men. These opposite effects of target sex and of informant sex might explain why no higher variances are found for males in self-reports of personality.

The Present Study

The present study serves three purposes: First, to test whether the findings in the study by Borkenau et al. (2013) hold also in culturally more diverse samples because – despite various differences – the four samples in that study were all European and thus stemmed from a similar cultural background. Second, to examine whether effects of target sex and of informant sex on the intrasex variability in personality differ between samples. Finally, if there are between-sample differences, to identify culture-level predictors which requires a sizable sample of cultures. In this context, it is interesting consulting cross-cultural studies on sex differences in *mean levels* of self-reported (Costa, Terracciano, & McCrae, 2001; Schmitt, Realo, Voracek, & Allik, 2008) and informant-reported (McCrae, Terracciano, & 78 members of the personality profiles of cultures project, 2005) personality traits: Surprisingly, the sex differences in mean levels were larger in the more developed and gender-egalitarian societies. Thus it seems worthwhile studying effects of indicators of the economic and social development of societies like the *Human Development Index* or the *Gender Inequality Index* (United Nations Development Programme, 2011) to investigate whether stronger sex differences are found in more developed societies not only for means but also for variances.

The data for the present analyses were collected in a large cross-cultural project on various aspects of informant reports of personality (McCrae, Terracciano, & 78 members of the personality profiles of cultures project, 2005; McCrae, Terracciano, & 79 members of the personality profiles of cultures project, 2005). In 51 cultures [1], a total of 12,156 college students participated. The 51 cultures included Belgium, the Czech Republic, Estonia, and

Germany, but the samples from these countries differed from those in the study by Borkenau et al. (2013). The 12,156 college students were asked to describe an individual from one of four target groups: college-aged men, college-aged women, adult men (> 40 years), or adult women (> 40 years). Raters were randomly assigned to one of the four target conditions. Thus distributions of target age and target sex were similar across cultures. Raters could choose as a target anyone they knew well, yielding a wide age and educational range. Although college students are certainly not representative of the general population, and even less so in less affluent countries, this approach made it feasible to obtain information on a wide range of targets in a wide range of cultures. More details on the data collection procedure are reported by McCrae, Terracciano, and 78 members of the personality profiles of cultures project (2005).

Nevertheless, the data quality varied considerably across cultures. Reasons were differences in the quality of the translation of the measurement instrument, that some items developed in Western societies seem to have no counterpart in some non-Western cultures, and that some samples were administered the inventory in a second language (e.g., French in Burkina Faso). Various indicators of data quality were therefore included and aggregated into a composite score that reflected the frequency of valid responses; lack of acquiescent responding or its opposite, nay-saying; number of missing responses (reversed); the participants' knowledge of the language in which the instrument was administered; and whether the translation of the measurement instrument had been published. More details on data quality as well as the languages in which the inventories were administered are reported by McCrae, Terracciano, and 78 members of the personality profiles of cultures project (2005).

Method

Participants

Translation and administration of the measurement instruments constituted an international collaborative effort. Valid peer reports were available for 12,156 target persons (51.9% female) from 48 countries, but as samples from French-speaking and German-speaking Switzerland, from England and Northern Ireland, and from China and Hong Kong were kept separate, the number of samples was 51. Sample sizes varied from $N = 106$ (Northern Ireland) to $N = 919$ (United States), 63.5% of the informants being female. The distribution of target age was bimodal with local maxima of 21 and 45 years, reflecting the instruction to describe a college-aged male or female person, or an adult male or female person beyond age 40.

Instrument

The measurement instrument was Form R of the Revised NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992), a 240-item measure of the Five-Factor model of personality. It measures the five basic personality dimensions Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C). Each of these personality domains comprises six facets. Thus there are 30 facets, each of them measured by eight items. Responses are given on five-point Likert scales with the endpoints *strongly disagree* and *strongly agree*. Form R was constructed for collecting observer ratings, its items being worded in the third person singular. Precautions taken to ensure the quality of the Form R translation are reported in more detail by McCrae, Terracciano, and 78 members of the personality profiles of cultures project (2005). Although the NEO PI-R does not

[1]In this article, the term *culture* is used loosely, referring to either nations or to subgroups within nations. We are aware that our samples do not necessarily reflect the full cultural diversity within nations.

include a scale measuring socially desirable responding, it does provide some checks for protocol validity that were used for excluding invalid protocols and for computing the index of data quality.

Culture-Level Variables

Five culture-level variables were included: First, the index of data quality as reported by McCrae, Terracciano, and 78 members of the personality profiles of cultures project (2005) for 50 cultures (all except Iran), and as reported by McCrae, Terracciano, and 79 members of the personality profiles of cultures project (2005) for Iran, was used as a control variable. Second, the Human Development Index (HDI) and, third, the Gender Inequality Index (GII) both being published annually by the United Nations (United Nations Development Programme, 2011). Running from 0 to 1, the HDI reflects life expectancy, years of schooling, and per-capita income, 1 indicating the highest development. It was included as an index of between-country differences in socio-economic development. The GII reflects female health variables (maternal mortality, adolescent fertility), male and female empowerment (percentage of females in the national parliament, percentages of males and females with at least secondary education), and female and male labor force participation. It also runs from 0 to 1, 1 indicating highest inequality, and was included as a measure of sex differences in behavioral options. These indices are reported for countries (although the HDI separately for China and Hong Kong), and so we used the same scores for the German-speaking and the French-speaking samples from Switzerland, and for the English and the Northern Ireland samples. Whereas the HDI was available for all countries, the GII was unavailable for Ethiopia, Hong Kong, Nigeria, and Serbia.

Finally we included Hofstede's (2001) measures of individualism (IND) and of masculinity (MASC), reflecting between-country differences in work goals. IND contrasts a focus on the self versus the family or group and was available for 40 cultures. MASC contrasts ego-directed work goals (earnings and advancements) with interpersonal relations goals (relationship with superiors, cooperation, friendly atmosphere) and was available for 38 of the 51 cultures under study.

Data Analyses

For descriptive purposes, we computed 1,530 variance ratios (VRs) separately for the 30 NEO PI-R facet scales and for the 51 samples, by dividing the intrasex variance between males by the intrasex variance between females, implying that VRs larger than 1.0 indicate more variability among men. These 1,530 VRs can be obtained from the first author. Before averaging VRs across facets or across samples, they were base-10 log transformed because VRs are not distributed normally, whereas their logarithms are approximately normally distributed (Hedges & Friedman, 1993). The averages were then back transformed, and these antilog VRs will be reported.

When computing VRs for the five personality domains (like in Table 1), we did *not* directly compute the VRs for the domain scores because these scores reflect the within-domain covariances between facets in addition to the facet variances. Rather, the VRs were calculated for each facet, their log transforms were averaged across facets within domains, and these averages were back transformed.

Multilevel Modeling

For statistical inference, we relied on hierarchical linear modeling (Raudenbush & Bryk, 2002; Raudenbush, Bryk, & Congdon, 2010), testing a three-level model. The outcome variable was the squared deviation of each target person's score from the culture-, sex-, and facet-specific mean. As a variance is the average squared deviation of individual

observations from their mean, the sample mean of squared deviation scores (SDS) is the variance in that sample.

Analyses within persons constituted Level 1. Here, we decomposed each participant's SDS for each facet into the participant's mean SDS across facets π_{0jk}, and the deviation e_{ijk} of the facet-specific SDS from that mean:

$$\mathrm{SDS}_{ijk}=\pi_{0jk}+e_{ijk} \quad (1)$$

i indicating the facet, *j* indicating the person, and *k* indicating the culture.

Level 2 represented differences between individual participants. Here we predicted the individual participants' average SDS (π_{0jk}) from the sex of the target and the sex of the informant:

$$\pi_{0jk}=\beta_{00k}+\beta_{01k} * \mathrm{Target\ \ Sex}_{jk}+\beta_{02k} * \mathrm{Informant\ \ Sex}_{jk}+r_{0jk} \quad (2)$$

Male had been coded 1 and *female* had been coded 2, implying that the coefficients β_{01k} and β_{02k} indicate the difference in the variances for female minus that for male targets and informants, respectively. Target and informant sex were entered grand-mean centered, thereby controlling for effects of different proportions of female targets and informants in the different cultures. Thus coefficient β_{00k} is the predicted average of all variances in culture *k*, assuming that the proportion of female targets and perceivers in that culture is representative of the proportions in the combined sample of 12,156 participants. In some analyses, we also entered the product of centered Target Sex and centered Informant Sex as a third predictor, to test for Target Sex × Informant Sex interactions (West, Ryu, Kwok, & Cham, 2011).

Level 3 represented the 51 cultures. Here, the coefficients in Formula 2 were predicted – first separately and then simultaneously – from the five culture-level variables Data Quality, HDI, GII, IND, and MASC. For example, the formula for predicting the effects of target sex simultaneously from the effects of data quality and the Human Development Index (a cross-level interaction) was:

$$\beta_{01k}=\gamma_{010}+\gamma_{011} * \mathrm{Data\ \ Quality}_{k}+\gamma_{012} * \mathrm{HDI}_{k}+u_{01k} \quad (3)$$

To facilitate comparisons between the effect sizes accounted for by the five culture-level predictors, these variables were standardized ($M = 0$; $SD = 1$). Thus, for example, γ_{012} in Formula 3 indicates the extent that, controlling for data quality, the difference between the variance among female targets minus the variance among male targets increases if HDI increases by one standard deviation.

Results

The VRs for target-sex are reported in Table 1 separately for the facets of N, E, O, A, and C. In addition, the rightmost column reports the VRs averaged across all 30 facets, and the bottom row reports the VRs averaged across cultures. Averaged across cultures, men varied more than women overall and on E, O, A, and C, but not on N. For five facets of N, more variance was found in descriptions of female targets, the exception being N2 (Angry Hostility; VR = 1.09). By contrast, men varied more than women on 22 of the altogether 24 facet scales of E, O, A, or C, the exceptions being E5 (Excitement Seeking; VR = 0.90), and A4 (Compliance; VR = 0.95).[2]

There were also systematic differences between cultures: Averaged across all 30 facets, the VRs varied from 0.85 for Japan to 1.34 for South Korea. The general trend, however, was that men varied more than women, 34 of the 51 row means exceeding 1.00, and 7 even exceeding 1.20. Moreover, in no culture did women vary more than men on all five dimensions, whereas in 14 cultures (Belgium, Czech Republic, England, Estonia, India, Lebanon, New Zealand, Poland, Portugal, Puerto Rico, South Korea, Slovak, Uganda, and the United States) men varied more than women on all five dimensions.[3]

Hierarchical Linear Modeling

First, we analyzed models with the Level-2 predictors target sex, informant sex, and their interaction (the product of centered target sex and centered informant sex), but without any culture-level predictors, allowing for random effects at all three levels. This was to check whether target sex, informant sex, and their interaction had significant effects on the outcome variable, and whether these effects varied between cultures. These analyses were run separately for N, E, O, A, and C with the SDS for their 6 facets as the outcome variable, and also for the combination of all 30 NEO PI-R facet scales with the SDS for all 30 facets as the outcome variable. As SDS are not distributed normally, robust standard errors were used.

Fixed effects of target and informant sex—The results for the Level-2 fixed effects are reported in Table 2. Unsurprisingly, the coefficients γ_{000} were all significant, implying significant variation in facet scores between persons. These coefficients indicate the average squared deviation of the individual observations from their culture-, sex-, and facet-specific means, that is, the averaged facet variances within the culture and gender groups.

The coefficients γ_{010}, indicating effects of target sex, were significant and negative for E, O, A, and C, as well as for all 30 facets combined, indicating more variance in descriptions of male than of female targets. In contrast, the coefficient for N was positive but not significant, suggesting similar variances in N for women and for men.

The coefficients γ_{020} that indicate effects of the sex of the informant were significant and positive in all analyses, implying that personality descriptions by female informants varied more than those by male informants. Finally, the coefficient γ_{030} reflecting systematic Target Sex × Informant Sex interactions, was significant for the facets of N only.

Differences between cultures in effects of target and informant sex—Next, we tested the Level-3 variance components for statistical significance to check whether the Level-2 effects varied significantly across cultures. If they did not, it would not be reasonable to search for any Level-3 predictors of differences between cultures. The findings are reported in Table 3. For the 30 facet scales combined, the intercept coefficients β_{00k} varied significantly, as did the coefficients β_{01k} for the effects of target sex, and β_{02k} for the effects of informant sex. Thus the cultures differed in their variances as well as the effects of target sex and informant sex on these variances. By contrast, the coefficients β_{03k}, representing Target Sex × Informant Sex interactions, did not vary significantly between cultures.

[2]We also computed VRs for domain scores by first adding the scores on the six relevant facets, and then dividing the variance in the sum score for men by the variance in the sum score for women. Averaged across cultures, this resulted in VRs of 0.97 (N), 1.07 (E), 1.13 (O), 1.07 (A), 1.16 (C), and 1.08 (average across the five domains). Thus this data-analytic approach did not result in substantially different findings.

[3]To check whether this effect was carried by the Belgian, Czech, Estonian, and German cultures having been studied by Borkenau et al. (2013) already, we repeated these analyses for the 47 remaining cultures. The average VRs then were 0.97 (N), 1.04 (E), 1.08 (O), 1.06 (A), and 1.11 (C), the grand mean being 1.05. Thus the findings were basically the same.

Analogous analyses of the differences between cultures were run separately for the facets of N, E, O, A, and C (Table 3). Throughout, the intercept coefficients β_{00k} varied significantly between cultures, whereas the effects of target sex (coefficients β_{01k}) varied for N, E, and A, but not for O or C, and the effects of informant sex (coefficients β_{02k}) varied significantly for E and C, but not for N, O, and A. Finally, none of the coefficients β_{03k}, representing Target Sex × Informant Sex interactions, varied significantly between cultures.

Culture-level variables as predictors of the effects of target sex and informant sex—As the mean of the Target Sex × Informant Sex interactions did not differ significantly from zero except for the N facets (Table 2), and these interactions did not differ significantly between cultures (Table 3), the Target Sex × Informant Sex interactions were dropped from all analyses predicting effects of target and informant sex from culture-level variables. Moreover, because the main effects of target sex and of informant sex varied significantly between cultures for the 30 facet scales combined, but for only some of the five personality dimensions analyzed separately (Table 3), the effects of the culture-level variables were tested for the combination of all 30 facet scales only, not separately for the facets of N, E, O, A, and C.

The correlations between the five culture-level predictors are reported in Table 4. With the exception of MASC, they were substantially correlated, indicating more individualism, more gender equality, and higher data quality in more developed societies.

The effects of the culture-level predictors were analyzed in three steps: First, each of them was entered separately in a series of five independent regression analyses to establish: (a) its total effect on the variance between participants within cultures (γ_{001}), (b) its cross-level interaction with the effects of target sex (γ_{011}), and (c) its cross-level interaction with the effects of informant sex (γ_{021}). The findings are reported in Table 5. The significant coefficients γ_{001} for each predictor except masculinity show that high data quality, a high HDI, a low GII, and high individualism predicted larger variances between persons: The variances in the personality traits rose or fell (in case of the GII) by about 10% (ratio of the coefficients γ_{001} reported In Table 5 and γ_{000} reported in Table 2) with a one standard deviation increase in these cultural variables. Similar findings have already been reported from the same dataset by McCrae, Terracciano, and 79 members of the personality profiles of cultures project (2005) who argued that the generally higher variances in rich, Western cultures might well reflect a higher data quality. We will come back to that later on.

Before, we would like to mention that high data quality, high HDI, low GII, and high IND were associated with stronger effects of target sex: The sex differences in variability in personality were more pronounced in the more developed, more gender-egalitarian, and more individualistic societies. For instance, the variances in the facet scale C4 (Achievement Striving) were 27.16 and 19.02 among US males and US females, respectively, implying a difference in the variances of 8.14, whereas in Morocco the corresponding variances were 14.58 and 13.97, that is, altogether lower and differing by 0.61 only. In contrast, effects of informant sex were not significantly associated with any of the five predictor variables, nor were any of the effects of MASC significant. Hofstede's masculinity measure was therefore not included in the further analyses.

Second, to clarify the unique contributions of the highly correlated culture-level variables, the predictors Data Quality, HDI, GII, and IND were simultaneously entered in a regression model predicting: (a) the variance between participants within cultures, (b) the effects of target sex on these variances, and (c) the effects of informant sex on these variances. High data quality and a high HDI predicted larger variances between participants, both ts > 2.56, ps < .02, but none of the four variables predicted stronger or weaker effects of target sex, all

*t*s < 1.74, *p*s > .09, or informant sex, all *t*s < 2.01, *p*s > .05, on these variances. This is likely to reflect the high correlations between these four culture-level predictors, leaving little unique variances that might be associated with effects of target sex or informant sex.

Finally, as data quality served mainly as a control variable, we reduced the number of culture-level predictors in the same analysis to two, entering data quality in combination with either HDI, GII, or IND. This resulted in three additional models. Results are reported in Table 6. High data quality predicted significantly larger variances in all three analyses, but not any effects of target sex or of informant sex on these variances. Moreover, a high HDI predicted larger variances when data quality was controlled. Finally and most interesting, individualism had marginally significant unique influences on the effects of target sex ($p = .052$; two-tailed test) and of perceiver sex ($p = .080$; two-tailed test): Higher variances in descriptions of male targets and by female perceivers occurred mostly in individualistic cultures.

Differences in Means and in Variances

That male targets varied more than female targets might reflect the fact that women had more extreme sample means than men on 19 of the 30 facet scales, that is, their scores deviated more strongly from a score of 16 on scales with a possible range of scores from 0 to 32. This occurred because the means for 24 of the 30 facets exceeded the scale midpoint of 16 in the coed sample, and women had higher means than men on 26 of the 30 facets. More extreme means, however, are associated with smaller variances (Wood & Wortman, 2012). Thus the sex differences in variances might be byproducts of sex differences in means. If that were the case, however, men should be overrepresented in only one tail of the distributions of facet scores, mostly the lower tail. By contrast, if men were overrepresented in both tails of the distributions, this would show that the higher variances for men are not merely artifacts of their less extreme means. Therefore, it was checked to what extent men and women were overrepresented in both tails of the distributions.

As such analyses make use of only a small fraction of the data, large sample sizes are needed. Therefore, we did not run these analyses separately for individual cultures. Moreover, we included only the 8,727 participants from those 34 cultures that showed averaged VRs > 1.00 (see Table 1) because, for the remaining 17 cultures, no higher variances between men had to be explained. Furthermore, to confirm the reliability of the findings, we ran separate analyses for the 4,093 (2,096 female) college-aged and the 4,634 (2,483 female) adult targets from the 34 cultures with a VR > 1.00. As female targets slightly outnumbered male targets in both samples, comparing the absolute frequency of men to the absolute frequency of women in the tails of each distribution would have been misleading. Rather, we chose a common cut-off score, and then counted the percentage of women and the percentage of men who exceeded that threshold. Specifically, we first identified the 5th and the 95th percentile in the coed sample, and then the percentage of men and the percentage of women with scores *more* extreme than these percentiles. For example, if 3% of the sample had scores exceeding 27, and 2% had a score of exactly 27, we counted which percentages of men and of women had scores higher than 27.

The findings are reported in Table 7. For six facets (college-aged sample) or seven facets (adult sample), a higher proportion of men than women was found in both tails. In the college-aged sample, this occurred for E1 (Warmth), O1 (Openness to Fantasy), O5 (Openness to Ideas), O6 (Openness to Values), A2 (Straightforwardness), and C1 (Competence). In the adult sample, that pattern was found for N2 (Angry Hostility), again O1, O5, and C1, and for C4 (Achievement Striving), C5 (Self-discipline), and C6 (Deliberation). By contrast, females were overrepresented in both tails of only one scale (N5, Impulsivity), and that in the college-aged sample only.[4]

We administered a binomial test to check whether the more frequent overrepresentation of males in both tails of the same distribution was significant, testing the null hypothesis that overrepresentation of females was as likely as overrepresentation of males. For the college-aged and the adult sample combined, the observed ratio of 13 : 1 was significant, $p <= .001$.[5] This illustrates nicely that, according to informant reports, extreme levels of personality traits occur more frequently among men than among women.

We also checked whether the higher variances in the descriptions by female than by male informants might reflect more extreme descriptions by men than by women. That, however, was not the case. Rather, for 25 of the 30 facet scales the means were more extreme for descriptions by female than by male informants. Higher variances in the descriptions by female judges did not reflect less extreme descriptions, accordingly.

Discussion

Consistent with the findings reported by Borkenau et al. (2013), the variances in the personality descriptions by informants were higher for male than for female targets. Moreover and also consistent with that study, descriptions by female informants varied more than descriptions by male informants. This pattern may explain why larger variances in descriptions of male than of female persons were found in informant reports and in observational studies (Mehl et al., 2007), but not in self-reports of personality (Borkenau et al., 2013; Cross et al., 2011).

It may seem at first glance that the overrepresentation of men in both tails of the distributions (Table 7) is too weak to justify such a conclusion, as that was found in only 13 out of 60 comparisons. But that apparently weak finding probably reflects that any sex difference in means operates against finding an overrepresentation of one sex in both tails of a distribution, even if there are substantial sex differences in variability. And sex differences in mean levels are a widespread phenomenon (McCrae, Terracciano, and 78 members of the personality profiles of cultures project, 2005). Thus if men are overrepresented in one tail of a distribution and women in the other tail reflecting different mean levels, that does not imply that there are no sex differences in variability.

Moreover, finding higher variances for men may seem inconsistent with the findings of a meta-analysis by Else-Quest, Hyde, Goldsmith, and van Hulle (2006) who analyzed sex differences in means and in variances in the temperament of children from 3 months to 13 years. Temperament was measured mostly by descriptions by informants (i.e., parents or teachers) in that study, but the variances were nevertheless not higher for boys than for girls. This suggests that the sex differences in variability found in the present study for college-aged persons and adults may not hold for children. This hypothesis is supported by Else-Quest et al.'s (2006) observation that "patterns of gender differences and similarities in temperament bear little resemblance to patterns of gender differences and similarities in adult personality" (p. 63). As the present study shows similar sex differences in variability in college-aged and in adult persons (Table 7), that differ from Else-Quest et al.'s (2006) findings for children, it seems reasonable that the sex differences in variability develop during puberty.

[4]We repeated these analyses for the unselected sample of 51 cultures. In the college-aged sample, men were then overrepresented in both tails of the distributions of four facets (E1, O1, O6, A2), and women were overrepresented in both tails of the distribution of one facet (N5). In the adult sample, men were overrepresented in both tails of the distributions of four facets (N2, C4, C5, C6), but women of no facet. Thus the effect was diluted but did not disappear when the analyses were run for an unselected sample of cultures.

[5]As the 30 NEO PI-R facets are correlated, not all assumptions of a binomial test are met. Thus the calculations of the α-error are not exact.

We also confirmed the result by Borkenau et al. (2013) that higher variances for male target persons are *not* found for neuroticism. Given that the samples in the two studies do not overlap, this suggests a reliable phenomenon that might reflect evolutionary, genetic, or social influences affecting only some personality traits, as none of these explanations implies that sex differences in variability should occur for all personality traits alike. Because of evolutionary pressure or gender-norms, males are more likely to suppress negative emotions (Brody, 2000; Anderson et al 2001 JPSP; Terracciano et al. 2003 JP), which may account for the contrasting pattern observed for N.

By contrast, the effects of informant sex seem to be quite similar for all personality domains: Descriptions by female informants vary more than descriptions by male informants. This may either reflect the fact that women vary more among each other in how they generally describe other persons, that is, they may show stronger perceiver effects as defined by Kenny (1994), or it may reflect that individual female informants differentiate more strongly than male informants between the trait levels of different target persons. Given that women seem to be more accurate judges of personality than men (Chan, Rogers, Parisotto, & Biesanz, 2011), it is likely that their personality judgments are more differentiated (Allik, Realo, Mõttus, Esko, Pullat, & Metspalu, 2010).

Cultural Differences

The intrasex variances as well as the effects of target and informant sex on these variances varied between cultures. Indeed, the effects of target sex were reversed in 15 of the 51 cultures under study. Here, women varied more in personality than men, although this reversal seems to reflect lower data quality to some extent. Such differences between cultures suggest that cultural factors are important, but may also reflect genetic factors as gene frequencies differ between world regions (Vogel & Motulsky, 1997). Genetic and cultural explanations are not mutually exclusive here, particularly as genetic and environmental influences both contribute to individual differences in personality (Borkenau, Riemann, Angleitner, & Spinath, 2001; Krueger & Johnson, 2008).

As already reported by McCrae, Terracciano, and 79 members of the personality profiles of cultures project (2005), the variances in the full samples were larger in more prosperous countries and those with a higher level of education, that is, those with a high HDI. Although this reflected higher data quality to some extent, it seems to hold also if data quality is controlled. Stronger personality differences in more developed societies may reflect more opportunities and more diverse behavioral options in such cultures, allowing a less restricted expression of personality dispositions, whereas expressions of personality may be more restricted in less developed societies reflecting lack of money, stricter social norms (Gelfand, Nishii, & Raver, 2006), lack of political freedom, and lack of educational and vocational opportunities.

Effects of Target Sex and Informant Sex

Interestingly, the effects of target sex were stronger in more individualistic societies. Further pursuing our explanation of larger personality differences in more developed societies, this might reflect that cultural individualism enhances the diversity of behavioral options for men more than for women. Assuming that women and men are still more involved in the family and in the vocational sphere respectively, and that behavioral options in the vocational sphere depend more strongly than those in the family sphere on the culture's individualism, it is reasonable that the behavioral options for women increase less with a culture's individualism than those for men. Therefore, high individualism may facilitate expressions of personality dispositions among men more than among women. In this context, it should be kept in mind that the present study is cross-sectional, studying effects

of differences between cultures. The longitudinal development within cultures in recent decades is another issue, its study requiring different data.

The effects of informant sex also varied significantly between cultures, but they were not associated with data quality, the HDI, and the GII. The effect of individualism, however, was marginally significant and did not shrink if data quality was controlled, indicating stronger effects of informant sex in more individualistic cultures. This may either reflect that women in more individualistic cultures vary more in how they describe other persons in general, for example in their leniency, or that individual female informants in such cultures differentiate more strongly between the trait levels of different persons.

Finally, only one measurement instrument, the NEO PI-R, was used in the present study, although in various translations. Thus future research should clarify whether the present findings generalize to other traits as well as to other instruments measuring the Big Five. Given the popularity of the Five-Factor model, however, and of the NEO PI-R as an instrument to measure its dimensions, the NEO PI-R is a good starting point.

Acknowledgments

Robert R. McCrae receives royalties from the Revised NEO Personality Inventory (NEO-PI-R). Czech participation was supported by Grant 406/01/1507 from the Grant Agency of the Czech Republic and is related to Research Plan AV 0Z7025918 of the Institute of Psychology, Academy of Sciences of the Czech Republic. Sami Gülgöz's participation was supported by the Turkish Academy of Sciences. Burkinabé and French Swiss participation was supported by a grant from the Swiss National Science Foundation to J. Rossier. The data collection in Hong Kong was supported by RGC Direct Allocation Grants DAG02/03.HSS14 and DAG03/04.HSS14 awarded to M. Yik. Data collection in Malaysia was supported by UKM Fundamental Research Grant 11JD/015/2003. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging.

We thank Ulrike Fritz and Olga Ostrovski for their help with the data analysis. The data were taken from a previous study (McCrae, Terracciano, and 79 members of the Personality Profiles of Cultures Project, 2005), and the contributions to that study by the international collaborators are gratefully acknowledged. In the order of data submission, these were: Brigitte Khoury, Florence Nansubuga, Goran Kneževič, Dragana Djuric Jocic, Hyun-nie Ahn, Chang-kyu Ahn, Filip De Fruyt, Sami Gülgöz, Willibald Ruch, M. Arif Ghayur, Maria D. Avia, Maria L. Sánchez- Bernardos, Jérôme Rossier, Donatien Dahourou, Ronald Fischer, Jane Shakespeare-Finch, Michelle Yik, Peter B. Smith, Alois Angleitner, Fritz Ostendorf, Magdalena S. Halim, Martina Hřebíčková, Thomas A. Martin, Tilahun Sineshaw, Andrzej Sekowski, Waldemar Klinkosz, Garry Prentice, Margaret McRorie, Carmen Flores-Mendoza, Yoshiko Shimonaka, Katsuharu Nakazato, Khairul A. Mastor, Claudio Barbaranelli, Lidia Alcalay, Franco Simonetti, V. S. Pramila, Ruth Falzon, Mary Ann Lauri, Mary Ann Borg Cunen, Sandra Scicluna Calleja, Margarida Pedroso de Lima, Denis Bratko, Iris Marušić, Jüri Allik, Anu Realo, Ahmed M. Abdel Khalek, Badr M. Alansari, Gregorio E. H. del Pilar, A. Oluyinka Ojedokun, Margaret Munyae, Louisa Budzinski, Shigehiro Oishi, Ed Diener, Niyada Chittcharat, Lei Wang, Andrew Beer, John Humrichouse, Erik Lykke Mortensen, Hans Henrik Jensen, Friðrik H. Jónsson, Emília Ficková, Lucia Adamovová, Velko S. Rus, Nuska Podobnik, Agata Zupancic, Rolando Diaz-Loving, Nora B. Leibovich, V. Schmidt, Norma Reátegui, Marina Brunner-Sciarra, Lindsay E. Ayearst, Krista K. Trobst, David Matsumoto, Aljoscha Neubauer, José Porrata, Jean-Pierre Rolland, Jean-Michel Petot, Nathalie Camart, Shahriar Shahidi.

References

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

Amos-Landgraf JM, Cottle A, Plenge RM, Friez M, Schwartz CE, Longshore J, Willard HF. X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. American Journal of Human Genetics. 2006; 79:493–499.10.1086/507565 [PubMed: 16909387]

Anderson C, John OP, Keltner D, Kring AM. Who attains social status? Effects of personality and physical attractiveness in social groups. Journal of Personality and Social Psychology. 2000; 81:116–132.10.1037/0022-3514.81.1.116 [PubMed: 11474718]

Archer J, Mehdikhani M. Variability among males in sexually selected attributes. Review of General Psychology. 2003; 7:219–236.10.1037/1089-2680.7.3.219

Arden R, Plomin R. Sex differences in variance of intelligence across childhood. Personality and Individual Differences. 2006; 41:39–48.10.1016/j.paid.2005.11.027

Benbow CP. Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. Behavioral and Brain Sciences. 1988; 11:169–183.

Borkenau P, Hřebíčková M, Kuppens P, Realo A, Allik J. Sex differences in variability in personality: A study in four samples. Journal of Personality. 2013; 81:51–62.10.1111/j.1467-6494.2012.00784.x

Borkenau P, Riemann R, Angleitner A, Spinath FM. Genetic and environmental influences on observed personality: Evidence from the German Observational Study of Adult Twins. Journal of Personality and Social Psychology. 2001; 80:655–668.10.1037//0022-3514.80.4.655 [PubMed: 11316228]

Bosak J, Sczesny S, Eagly AH. The impact of social roles on trait judgments: A critical reexamination. Personality and Social Psychology Bulletin. 2012; 38:429–440.10.1177/0146167211427308 [PubMed: 22201645]

Brody, LR. The socialization of gender differences in emotional expression: Display rules, infant temperament, and differentiation. In: Fischer, AH., editor. Gender and emotion: Social psychological perspectives. New York: Cambridge University Press; 2000. p. 24-47.

Chan M, Rogers KH, Parisotto KL, Biesanz JC. Forming first impressions: The role of gender and normative accuracy in personality perception. Journal of Research in Personality. 2011; 45:117–120.10.1016/j.jrp.2010.11.001

Check E. Genetics: The X factor. Nature. 2005; 434:266–267.10.1038/434266a [PubMed: 15772620]

Costa, PT.; McCrae, RR. NEO PI-R professional manual. Orlando, FL: Psychological Assessment Resources; 1992.

Costa PT, Terracciano A, McCrae RR. Gender differences in personality traits across cultures: Robust and surprising findings. Journal of Personality and Social Psychology. 2001; 81:322–331.10.1037//0022-3514.81.2.322 [PubMed: 11519935]

Cross CP, Copping LT, Campbell A. Sex differences in impulsivity: A meta-analysis. Psychological Bulletin. 2011; 137:97–130.10.1037/a0021591 [PubMed: 21219058]

Deary IJ, Thorpe G, Wilson V, Starr JM, Whalley LJ. Population sex differences in IQ at the age 11: The Scottish mental survey 1932. Intelligence. 2003; 31:533–542.10.1016/S0160-2896(03)00053-9

Eagly AH, Steffen VJ. Gender and aggressive behavior: A meta-analytic review of the social psychological literature. Psychological Bulletin. 1986; 100:309–330.10.1037//0033-2909.100.3.309 [PubMed: 3797558]

Else-Quest NM, Hyde JS, Goldsmith HH, Van Hulle CA. Gender differences in temperament: A meta-analysis. Psychological Bulletin. 2006; 132:33–72.10.1037/0033-2909.132.1.3 [PubMed: 16435957]

Feingold A. The additive effects of differences in central tendency and variability are important in comparisons between groups. American Psychologist. 1995; 50:5–13.10.1037//0003-066X.50.1.5

Gelfand MJ, Nishii LH, Raver JL. On the nature and importance of cultural tightness-looseness. Journal of Applied Psychology. 2006; 91:1225–1244.10.1037/0021-9010.91.6.1225 [PubMed: 17100480]

Hedges LV, Friedman L. Gender differences in variability in intellectual abilities: A reanalysis of Feingold's results. Review of Educational Research. 1993; 63:94–105.10.3102/00346543063001094

Hedges LV, Nowell A. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. Science. 1995; 269:41–45.10.1126/science.7604277 [PubMed: 7604277]

Hofstede, G. Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations. 2. Thousand Oaks, CA: Sage; 2001.

Humphreys LG. Sex differences in variability may be more important than sex differences in means. Behavioral and Brain Sciences. 1988; 11:195–196.

Kenny, DA. Interpersonal perception: A social relations analysis. New York: Guilford Press; 1994.

Krueger, RF.; Johnson, W. Behavioral genetics and personality: A new look at the integration of nature and nurture. In: John, OP.; Robins, RW.; Pervin, LA., editors. Handbook of personality: Theory and research. 3. NY: Guilford Press; 2008. p. 287-310.

Lehre AC, Lehre KP, Laake P, Danbolt NC. Greater intrasex phenotype variability in males than in females is a fundamental aspect of the gender differences in humans. Developmental Psychobiology. 2009; 51:198–206.10.1002/dev.20358 [PubMed: 19031491]

McCrae, RR.; Costa, PT, Jr. NEO Inventories professional manual. Odessa, FL: Psychological Assessment Resources; 2010.

McCrae RR, Terracciano A. 78 members of the personality profiles of cultures project. Universal features of personality traits from the observer's perspective: Data from 50 cultures. Journal of Personality and Social Psychology. 2005; 88:547–561.10.1037/0022-3514.88.3.547 [PubMed: 15740445]

McCrae RR, Terracciano A. 79 members of the personality profiles of cultures project. Personality profiles of cultures: Aggregate personality traits. Journal of Personality and Social Psychology. 2005; 89:407–425.10.1037/0022-3514.89.3.407 [PubMed: 16248722]

Mehl MR, Vazire S, Ramírez-Esparza N, Slatcher RB, Pennebaker JW. Are women really more talkative than men? Science. 2007; 317:82.10.1126/science.1139940 [PubMed: 17615349]

Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2. Thousand Oaks, CA: Sage; 2002.

Raudenbush, SW.; Bryk, AS.; Congdon, R. HLM7: Hierarchical linear and nonlinear modelling [Computer software]. Lincolnwood, IL: Scientific Software International; 2010.

Schmitt DP, Realo A, Voracek M, Allik J. 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. 2008; 94:168–182.10.1037/0022-3514.94.1.168 [PubMed: 18179326]

Terracciano A, McCrae RR, Hagemann D, Costa PT. Individual difference variables, affective differentiation, and the structures of affect. Journal of Personality. 2003; 71:669–703.10.1111/1467-6494.7105001 [PubMed: 12932207]

United Nations Development Programme. Human development report 2011: Human development statistical annex. 2011. Retrieved from http://hdr.undp.org/en/media/HDR_2011_EN_Tables.pdf

Vogel, F.; Motulsky, AG. Human genetics: Problems and approaches. Berlin (Germany): Springer; 1997.

West SG, Ryu E, Kwok O, Cham H. Multilevel modeling: Current and future applications in personality research. Journal of Personality. 2011; 79:2–50.10.1111/j.1467-6494.2010.00681.x [PubMed: 21223263]

Wood D, Wortman J. Trait means and desirabilities as artifactual and real sources of differential stability of personality traits. Journal of Personality. 2012; 80:665–701.10.1111/j.1467-6494.2011.00740.x [PubMed: 22092250]

Highlights

- We studied effects of sex on the variances in personality descriptions.
- Informant reports of personality varied more for male than for female targets.
- Descriptions by female informants varied more than descriptions by male informants.
- Across 51 cultures, both effects were stronger in more individualistic societies.

Table 1

Variance Ratios for Target Sex

Culture	N Facets	E Facets	O Facets	A Facets	C Facets	Mean
Argentina	0.90	1.19	0.93	1.43	1.14	1.10
Australia	0.94	0.88	1.07	1.18	0.97	1.00
Austria	0.99	1.12	0.98	1.06	1.11	1.05
Belgium	1.01	1.05	1.13	1.19	1.10	1.09
Botswana	0.99	0.99	1.01	0.95	0.92	0.97
Brazil	0.95	1.04	1.12	1.02	1.07	1.04
Burkina Faso	0.80	0.92	0.95	1.18	1.07	0.98
Canada	1.01	0.88	1.05	1.16	1.08	1.03
Chile	0.87	1.12	1.14	1.03	1.16	1.06
China	1.12	0.98	1.21	1.01	1.33	1.12
Croatia	0.89	0.96	1.09	0.98	1.24	1.03
Czech	1.07	1.34	1.23	1.39	1.35	1.27
Denmark	0.90	1.15	1.49	0.89	1.23	1.11
England	1.06	1.45	1.02	1.43	1.29	1.24
Estonia	1.03	1.10	1.10	1.17	1.32	1.14
Ethiopia	1.12	0.97	1.02	0.88	1.24	1.04
France	0.85	1.04	1.16	0.93	1.16	1.02
Germany	0.95	1.28	1.40	1.19	1.26	1.21
Hong Kong	1.05	0.95	1.02	0.88	1.02	0.98
Iceland	0.93	1.11	0.99	1.02	1.17	1.04
India	1.23	1.10	1.16	1.07	1.23	1.15
Indonesia	0.91	0.95	1.00	0.96	1.06	0.97
Iran	0.87	0.87	0.83	1.14	1.12	0.95
Italy	0.79	1.00	1.27	1.01	0.83	0.97
Japan	0.76	0.85	1.01	0.73	0.92	0.85
Kuwait	0.89	1.09	1.28	0.95	1.09	1.05
Lebanon	1.03	1.04	1.04	1.05	1.12	1.05
Malaysia	1.03	0.98	1.02	1.25	1.19	1.09
Malta	0.78	0.97	1.03	0.84	1.04	0.93

Culture	N Facets	E Facets	O Facets	A Facets	C Facets	Mean
Mexico	0.99	0.89	0.95	1.15	0.90	0.97
Morocco	1.00	0.95	1.20	1.02	1.18	1.07
North Ireland	1.09	1.24	1.24	0.99	1.13	1.13
New Zealand	1.23	1.31	1.21	1.19	1.12	1.21
Nigeria	0.97	0.96	0.98	1.00	1.01	0.99
Peru	0.74	0.93	1.03	0.72	0.96	0.87
Philippines	0.89	0.85	0.97	0.91	1.09	0.94
Poland	1.07	1.19	1.07	1.38	1.07	1.15
Portugal	1.31	1.00	1.12	1.31	1.31	1.20
Puerto Rico	1.38	1.29	1.37	1.12	1.26	1.28
Russia	0.94	0.96	0.98	0.99	1.01	0.97
Serbia	0.81	0.95	1.16	1.05	0.92	0.97
Slovak	1.03	1.16	1.27	1.12	1.24	1.16
Slovenia	0.95	0.95	0.86	1.17	1.24	1.03
South Korea	1.25	1.35	1.27	1.47	1.36	1.34
Spain	0.90	1.31	1.30	1.31	1.14	1.18
Swiss-French	0.86	0.84	1.06	1.11	1.08	0.98
Swiss-German	0.96	1.00	1.22	1.06	1.01	1.05
Thailand	0.92	0.98	0.91	1.14	1.08	1.00
Turkey	1.01	0.98	1.15	0.97	0.99	1.02
Uganda	1.25	1.21	1.00	1.14	1.37	1.19
USA	1.05	1.21	1.14	1.08	1.20	1.13
Mean	0.98	1.05	1.09	1.07	1.12	1.06

Note: The entries are the ratios of the variances among male divided by the variances among female target persons. N = Neuroticism; E = Extraversion; O = Openness to Experience; A = Agreeableness; C = Conscientiousness.

Table 2

Main and Interactive Effects of the Level-2 Predictors Target Sex and Informant Sex.

	Coefficient	Standard error [a]	*t*-ratio	*p*-value
N Facets				
Intercept, γ_{000}	23.95	0.73	32.89	<.001
Target sex, γ_{010}	0.48	0.43	1.10	.275
Informant sex, γ_{020}	2.24	0.45	4.96	<.001
Interaction, γ_{030}	-1.45	0.70	-2.08	.043
E Facets				
Intercept, γ_{000}	25.44	0.71	35.65	<.001
Target sex, γ_{010}	-1.43	0.55	-2.60	.012
Informant sex, γ_{020}	3.23	0.49	6.58	<.001
Interaction, γ_{030}	-0.59	0.90	-0.65	.517
O Facets				
Intercept, γ_{000}	22.70	0.82	27.70	<.001
Target sex, γ_{010}	-2.61	0.43	-6.12	<.001
Informant sex, γ_{020}	2.26	0.42	5.45	<.001
Interaction, γ_{030}	0.66	0.63	1.05	.299
A Facets				
Intercept, γ_{000}	24.92	0.68	36.39	<.001
Target sex, γ_{010}	-1.97	0.53	-3.69	<.001
Informant sex, γ_{020}	2.53	0.47	5.38	<.001
Interaction, γ_{030}	-1.24	0.94	-1.32	.194
C Facets				
Intercept, γ_{000}	26.99	0.66	40.89	<.001
Target sex, γ_{010}	-3.31	0.47	-7.04	<.001
Informant sex, γ_{020}	3.08	0.62	4.95	<.001
Interaction, γ_{030}	0.54	1.04	0.52	.621
All 30 facets combined				
Intercept, γ_{000}	24.80	0.67	36.88	<.001
Targec sex, γ_{010}	-1.73	0.36	-4.78	<.001
Informant sex, γ_{020}	2.62	0.37	7.00	<.001
Interaction, γ_{030}	-0.39	0.59	-0.67	.505

Note:

[a] Robust standard errors.

Table 3

Variances Across Cultures of Level-2 Parameter Estimates

	Variance	χ^2	*p*-value
N Facets			
Intercept, β_{00k}	25.35	567.96	<.001
Target sex, β_{01k}	2.22	74.35	.009
Informant sex, β_{02k}	2.49	57.21	.170
Interaction, β_{03k}	1.34	41.79	>.500
E Facets			
Intercept, β_{00k}	23.82	607.88	<.001
Target sex, β_{01k}	7.14	93.01	<.001
Informant sex, β_{02k}	2.87	66.41	.049
Interaction, β_{03k}	4.55	53.68	.299
O Facets			
Intercept, β_{00k}	32.49	939.05	<.001
Target sex, β_{01k}	2.96	65.04	.075
Informant sex, β_{02k}	1.78	59.43	.170
Interaction, β_{03k}	1.04	40.08	>.500
A Facets			
Intercept, β_{00k}	21.25	409.56	<.001
Target sex, β_{01k}	5.15	75.81	.008
Informant sex, β_{02k}	2.84	55.69	.237
Interaction, β_{03k}	8.48	53.60	.302
C Facets			
Intercept, β_{00k}	18.73	337.03	<.001
Target sex, β_{01k}	1.11	48.02	>.500
Informant sex, β_{02k}	5.98	72.80	.019
Interaction, β_{03k}	6.16	46.94	>0.500
All 30 facets combined			
Intercept, β_{00k}	21.97	1032.03	<.001
Target sex, β_{01k}	2.56	88.08	<.001
Informant sex, β_{02k}	2.42	74.95	.010
Interaction, β_{03k}	0.74	45.45	>.500

Note: Interaction = Target Sex × Perceiver Sex Interaction. With 51 cultures, there were 50 degrees of freedom.

Table 4

Correlations Among Culture-Level Predictors.

	HDI	GII	IND	MASC
Data Quality	.63**	-.67**	.50**	.09
HDI		-.86**	.58**	.22
GII			-.55**	-.10
Individualism				.41*

Note: HDI = Human Development Index; GII = Gender Inequality Index; IND = Hofstede's measure of individualism; MASC = Hofstede's measure of masculinity.

* $p < .05$.

** $p < .01$.

Table 5

Total Effects of Level-3 Predictors on the Variance Between Persons (γ_{001}), and Their Cross-Level Interactions with Target Sex (γ_{011}) and Informant Sex (γ_{021}).

	Coefficient	Standard error [a]	*t*-ratio	*p*-value
Data Quality				
Intercept, γ_{001}	3.29	0.62	5.32	<.001
Target sex, γ_{011}	-0.80	0.31	-2.55	.014
Informant sex, γ_{021}	-0.23	0.37	-0.61	.547
Human Development Index				
Intercept, γ_{001}	3.11	0.44	7.03	<.001
Target sex, γ_{011}	-0.85	0.29	-2.93	.005
Informant sex, γ_{021}	0.36	0.34	1.06	.296
Gender Inequality Index				
Intercept, γ_{001}	-2.21	0.57	-3.88	<.001
Target sex, γ_{011}	0.85	0.32	2.64	.011
Informant sex, γ_{021}	0.24	0.44	0.53	.598
Hofstede's Measure of Individualism				
Intercept, γ_{001}	2.05	0.52	3.95	<.001
Target sex, γ_{011}	-0.98	0.33	-3.00	.005
Informant sex, γ_{021}	0.60	0.41	1.45	.155
Hofstede Measure of Masculinity				
Intercept, γ_{011}	-0.24	0.55	-0.43	.669
Target sex, γ_{011}	0.46	0.47	0.96	.341
Informant sex, γ_{021}	-0.57	0.42	-1.38	.177

Note:

[a] Robust standard errors.

J Res Pers. Author manuscript; available in PMC 2014 April 01.

Table 6

Unique Effects of data quality, HDI, GII, and Individualism on the Variances Between Persons, and on the Effects of Target Sex and Informant Sex on these Variances.

	Model 1			Model 2			Model 3		
	Coefficient	***t*[a]**	***p***	**Coefficient**	***t*[a]**	***p***	**Coefficient**	***t*[a]**	***p***
Intercept									
Data Quality, γ_{001}	2.18	3.04	.004	2.68	3.40	.001	2.97	4.60	<.001
HDI, γ_{002}	1.78	3.89	<.001						
GII, γ_{002}				-0.45	-0.67	.507			
IND, γ_{002}							0.58	0.88	.382
Target Sex									
Data Quality, γ_{011}	-0.47	-1.19	.238	-0.51	-1.32	.194	-0.45	-1.13	.265
HDI, γ_{012}	-0.52	-1.43	.159						
GII, γ_{012}				0.50	1.31	.198			
IND. γ_{012}							-0.76	-2.01	.052
Informant Sex									
Data Quality, γ_{021}	-0.65	-1.39	.172	-0.23	-0.44	.661	-0.59	-1.19	.244
HDI, γ_{022}	0.74	1.75	.087						
GII, γ_{022}				.107	.181	.857			
IND, γ_{022}							0.85	1.80	.080

Note: HDI = Human Development Index. GII = Gender Inequality Index. IND = Hofstede's measure of individualism.

[a] t-values are based on robust standard errors.

Table 7

Percentages of College-Aged and Adult Women and Men in 34 Cultures Below the 5th or Above the 95th Percentile.

	College age				Adult age			
	< 5th Percentile		> 95th Percentile		< 5th Percentile		> 95th Percentile	
	women	men	women	men	women	men	women	men
N1	2.4	6.9	4.7	2.1	2.0	5.5	4.9	1.5
N2	2.7	4.7	4.2	3.9	**4.4**	**5.5**	**3.7**	**5.7**
N3	2.1	4.6	4.5	3.3	2.9	5.4	4.8	2.6
N4	1.8	5.2	4.6	2.3	2.2	4.7	4.1	2.2
N5	*4.2*	*4.1*	*4.8*	*4.3*	5.1	3.7	3.5	4.9
N6	2.7	5.8	5.4	3.7	1.9	5.0	5.2	3.3
E1	**2.9**	**5.3**	**4.4**	**5.1**	2.6	6.4	4.8	4.2
E2	3.5	6.4	5.2	3.7	2.5	6.1	4.6	3.1
E3	4.7	4.2	3.5	4.1	5.1	2.9	2.8	6.0
E4	2.7	4.0	5.8	4.0	2.4	4.8	3.8	3.6
E5	4.6	2.5	3.9	5.4	5.7	2.6	2.5	5.2
E6	2.6	5.2	5.6	2.7	2.1	5.4	4.7	4.3
O1	**2.9**	**4.1**	**3.8**	**4.6**	**3.3**	**6.0**	**4.9**	**5.0**
O2	2.6	6.5	4.2	2.4	2.4	7.3	5.6	3.1
O3	3.1	5.7	5.8	3.5	2.3	7.5	5.8	1.7
O4	2.9	5.4	3.7	3.1	2.4	5.0	3.9	2.6
O5	**3.3**	**3.5**	**2.7**	**6.2**	**4.3**	**11.9**	**2.7**	**6.1**
O6	**4.1**	**4.7**	**3.6**	**4.7**	2.8	3.5	4.4	2.3
A1	2.9	4.9	4.3	4.2	2.7	4.2	4.1	3.2
A2	**3.4**	**4.6**	**3.8**	**4.0**	3.5	6.2	5.0	3.6
A3	3.3	5.1	3.9	3.5	2.6	6.1	4.6	2.2
A4	3.6	3.7	4.9	4.0	3.1	5.4	5.0	3.0
A5	3.2	5.2	4.7	4.5	2.4	5.5	4.7	3.2
A6	3.1	5.4	3.4	2.6	2.0	5.3	5.0	1.9
C1	**4.2**	**4.9**	**3.4**	**3.6**	**3.9**	**4.1**	**1.9**	**3.7**
C2	2.8	4.2	5.5	3.9	2.6	6.6	3.4	2.4
C3	2.6	5.7	4.4	3.5	2.6	5.1	3.6	3.6

	College age				Adult age			
	< 5th Percentile		> 95th Percentile		< 5th Percentile		> 95th Percentile	
	women	men	women	men	women	men	women	men
C4	2.5	5.1	4.3	3.5	**3.6**	**4.2**	**2.6**	**5.2**
C5	3.0	5.3	5.1	3.3	**3.0**	**5.5**	**4.5**	**5.1**
C6	2.9	4.2	4.6	4.3	**4.3**	**5.2**	**3.3**	**5.7**

Note: N1 = Anxiety, N2 = Angry hostility, N3 = Depression, N4 = Self-consciousness, N5 = Impulsiveness, N6 = Vulnerability, E1 = Warmth, E2 = Gregariousness, E3 = Assertiveness, E4 = Activity, E5 = Excitement seeking, E6 = Positive emotions, O1 = Fantasy, O2 = Aesthetics, O3 = Feelings, O4 = Actions, O5 = Ideas, O6 = Values, A1 = Trust, A2 = Straightforwardness, A3 = Altruism, A4 = Compliance, A5 = Modesty, A6 = Tender-Mindedness, C1 = Competence, C2 = Order, C3 = Dutifulness, C4 = Achievement striving, C5 = Self-discipline, C6 = Deliberation.
Bold font indicates that men are overrepresented in both tails of a distribution. Italics indicate that women are overrepresented in both tails of a distribution.