Gender differences in personality across the ten aspects of the Big Five

Yanna J. Weisberg[1], Colin G. DeYoung[2]* and Jacob B. Hirsh[3]

[1] Department of Psychology, Linfield College, McMinnville, OR, USA
[2] Department of Psychology, University of Minnesota, Minneapolis, MN, USA
[3] University of Toronto, Toronto, ON, Canada

This paper investigates gender differences in personality traits, both at the level of the Big Five and at the sublevel of two aspects within each Big Five domain. Replicating previous findings, women reported higher Big Five Extraversion, Agreeableness, and Neuroticism scores than men. However, more extensive gender differences were found at the level of the aspects, with significant gender differences appearing in both aspects of every Big Five trait. For Extraversion, Openness, and Conscientiousness, the gender differences were found to diverge at the aspect level, rendering them either small or undetectable at the Big Five level. These findings clarify the nature of gender differences in personality and highlight the utility of measuring personality at the aspect level.

Keywords: personality, Big Five, gender differences

INTRODUCTION

Men and women belong to different species and communications between them is still in its infancy. – Bill Cosby

Many people, including Bill Cosby, perceive the differences between men and women to be large – so large, in fact, that communication between genders may be difficult. Countless examples from popular culture reinforce this view of extreme differences between the sexes – but is it accurate? Men and women have obviously different biological roles when it comes to propagation of the species, but how much they differ psychologically is a more controversial question, one that requires empirical research to answer adequately. Whether the underlying causes of psychological gender differences are evolutionary or socio-cultural, understanding how men and women differ in the ways in which they think, feel, and behave can shed light on the human condition.

The study of personality is particularly useful in attempting to examine psychological differences between genders. Personality is often conceptualized as the extent to which someone displays high or low levels of specific traits. Traits are the consistent patterns of thoughts, feelings, motives, and behaviors that a person exhibits across situations (Fleeson and Gallagher, 2009). That is, someone who scores high on a trait will exhibit psychological states related to that trait more often and to a greater extent than individuals who score low on that trait.

Gender differences in personality traits are often characterized in terms of which gender has higher scores on that trait, on average. For example, women are often found to be more agreeable than men (Feingold, 1994; Costa et al., 2001). This means that women, on average, are more nurturing, tender-minded, and altruistic more often and to a greater extent than men. However, such a finding does not preclude the fact that men may also experience nurturing, tender-minded, and altruistic states, and that some men may even score higher in these traits than some women. The goal of investigating gender differences in personality, therefore, is to elucidate the differences among general patterns of behavior in men and women on average, with the understanding that both men and women can experience states across the full range of most traits. Gender differences in terms of mean differences do not imply that men and women only experience states on opposing ends of the trait spectrum; on the contrary, significant differences can exist along with a high degree of overlap between the distributions of men and women (Hyde, 2005).

A core mission of personality psychology has been the development of an adequate taxonomy of personality traits. Drawing on trait descriptors used in natural language (selected from dictionaries) and in personality questionnaires, a five factor structure has emerged to explain covariation among traits. The five factor model or *Big Five* categorizes traits into the broad domains of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness/Intellect (Digman, 1990; John et al., 2008).

Gender differences in personality are often examined in terms of the Big Five. However, the Big Five do not exhaust all of the important distinctions among personality traits. Traits are hierarchically organized such that more specific traits that vary together are grouped within higher-order factors, like the Big Five. In the study of gender differences, therefore, one can investigate gender differences in personality traits at multiple levels of resolution. Most trait research has focused on two levels of traits: (1) the broad Big Five domains and (2) many more specific traits, called facets, which are grouped together within the Big Five. Currently, there is no consensus as to the identity and number of facets within the Big Five. Different approaches have identified different sets of facets, based on rational review of psychological constructs

(e.g., Costa and McCrae, 1992) or by systematic sampling from the space defined by pairs of Big Five factors (e.g., Soto and John, 2009). In the present study, we utilized an empirically identified level of personality traits that falls between narrow facets and broad domains. This level of personality organization has the potential to characterize gender differences with a finer grain of detail than the Big Five, revealing differences that are obscured in the Big Five. Additionally, it provides an empirically based taxonomy of lower-level traits, that is more likely to represent an adequate taxonomy of traits than existing facet models.

If the Big Five constituted the level of the personality hierarchy immediately above the facets, only one factor should be necessary to explain the shared variance of the facets within a given Big Five domain. However, a large behavioral genetic study revealed that *two* distinct factors were necessary to account for the shared genetic variance among the facets within each domain (Jang et al., 2002). In a separate study using factor analysis of 15 different facets within each domain, two phenotypic factors similar to the genetic factors were found for each of the Big Five dimensions (DeYoung et al., 2007). This research indicates that each of the Big Five contains two separable, though correlated, aspects, reflecting a level of personality below the broad domains but above the many facet scales. DeYoung et al. (2007) characterized these aspects by examining their factor-score correlations with over 2000 items from the International Personality Item Pool (IPIP). The aspects were labeled as follows: Volatility and Withdrawal for Neuroticism; Enthusiasm and Assertiveness for Extraversion; Intellect and Openness for Openness/Intellect; Industriousness and Orderliness for Conscientiousness; and Compassion and Politeness for Agreeableness. The aspect level of traits may be especially useful for the investigation of gender differences because these differences are sometimes unclear at the Big Five level and can be large and in opposite directions at the facet level. The aspects provide a non-arbitrary and parsimonious system for examining gender differences at a level of traits more specific than the Big Five.

Gender differences have been documented for a number of personality traits. Most meta-analyses and reviews examine gender differences in self-reports of personality on questionnaires that measure the Big Five, as well as facets within each (Feingold, 1994; Costa et al., 2001; Lippa, 2010). To our knowledge, however, no analyses have specifically examined the two aspects of each Big Five trait.

GENDER DIFFERENCES IN BIG FIVE PERSONALITY TRAITS

The investigation of personality differences is important to our understanding of general human variation, though it is not without controversy. Research on individual differences in intelligence, for example, has sparked years of scientifically and emotionally motivated debate (Neisser et al., 1996). Gender differences research has also proven to be controversial, with much of the debate concerning the causes and precursors of differences. Biological and evolutionary approaches posit that gender differences are due to men and women's dimorphically evolved concerns with respect to reproductive issues, parental investment in offspring (Trivers, 1972; Buss, 2008). According to these theories,

women should be more concerned with successfully raising children and should therefore be more cautious, agreeable, nurturing, and emotionally involved. Men, on the other hand, should be more concerned with obtaining viable mating opportunities and should therefore exhibit more Assertiveness, risk-taking, and aggression. Other theories suggest that gender norms are shaped by socio-cultural influences, such that women and men are expected to serve different roles in society and are therefore socialized to behave differently from one another (Wood and Eagly, 2002; Eagly and Wood, 2005). Of course, it may well be that both evolutionary and social forces have contributed to gender differences. Interestingly, recent studies have shown that gender differences in personality tend to be larger in more developed, Western cultures with less traditional sex roles (Costa et al., 2001; Schmitt et al., 2008). In our review, we focus on the patterns that have been found most consistently across cultures. The overall pattern for gender differences in personality measured by the Big Five is that existing differences are small to medium in size. For some domains, the gender differences are in the same direction across all measured facets; for others, however, the patterns are more divergent.

Neuroticism

Neuroticism describes the tendency to experience negative emotion and related processes in response to perceived threat and punishment; these include anxiety, depression, anger, self-consciousness, and emotional lability. Women have been found to score higher than men on Neuroticism as measured at the Big Five trait level, as well as on most facets of Neuroticism included in a common measure of the Big Five, the NEO-PI-R (Costa et al., 2001). Additionally, women also score higher than men on related measures not designed specifically to measure the Big Five, such as indices of anxiety (Feingold, 1994) and low self-esteem (Kling et al., 1999). The one facet of Neuroticism in which women do not always exhibit higher scores than men is Anger, or Angry Hostility (Costa et al., 2001).

Agreeableness

Agreeableness comprises traits relating to altruism, such as empathy and kindness. Agreeableness involves the tendency toward cooperation, maintenance of social harmony, and consideration of the concerns of others (as opposed to exploitation or victimization of others). Women consistently score higher than men on Agreeableness and related measures, such as tender-mindedness (Feingold, 1994; Costa et al., 2001).

Conscientiousness

Conscientiousness describes traits related to self-discipline, organization, and the control of impulses, and appears to reflect the ability to exert self-control in order to follow rules or maintain goal pursuit. Women score somewhat higher than men on some facets of Conscientiousness, such as order, dutifulness, and self-discipline (Feingold, 1994; Costa et al., 2001). These differences, however, are not consistent across cultures, and no significant gender difference has typically been found in Conscientiousness at the Big Five trait level (Costa et al., 2001).

Extraversion

Extraversion reflects sociability, Assertiveness, and positive emotionality, all of which have been linked to sensitivity to rewards (Depue and Collins, 1999; DeYoung and Gray, 2009). Whereas gender differences are small on the overall domain level of Extraversion (with women typically scoring higher), the small effect size could be due to the existence of gender differences in different directions at the facet level. Women tend to score higher than men on Warmth, Gregariousness, and Positive Emotions, whereas men score higher than women on Assertiveness and Excitement Seeking (Feingold, 1994; Costa et al., 2001).

Extraversion, together with Agreeableness, can be used to describe the two dimensions of the interpersonal circumplex (IPC; Wiggins, 1979), which contains descriptions of traits relevant to interpersonal interaction. Though originally posited to describe interpersonal traits using axes of Love and Status/Dominance, the IPC can also be conceptualized as a rotation of Big Five Extraversion and Agreeableness (McCrae and Costa, 1989). Given the importance of Extraversion to the interpersonal domain, it may be expected that women would consistently score higher than men. However the pole of the IPC often called Dominance contains traits such as bossy, domineering, and assertive. Men tend to be more dominant and agentic than women, and exhibit higher levels of these traits (Helgeson and Fritz, 1999). Gender differences in Extraversion may therefore switch directions depending on whether the specific traits measured fall closer or further from the dominance pole.

Openness/Intellect

Openness/Intellect reflects imagination, creativity, intellectual curiosity, and appreciation of esthetic experiences. Broadly, Openness/Intellect relates to the ability and interest in attending to and processing complex stimuli. No significant gender differences are typically found on Openness/Intellect at the domain level, likely due to the divergent content of the trait. For example, women have been found to score higher than men on the facets of Esthetics and Feelings (Costa et al., 2001), whereas men tend to score higher on the Ideas facet (Feingold, 1994; Costa et al., 2001).

HYPOTHESES REGARDING THE 10 ASPECTS

The pattern of gender differences reviewed above highlights the need to look beyond the Big Five level to traits at lower levels of analysis. Because the domains of the Big Five are so broad and encompass a variety of personality characteristics, greater specificity is needed to uncover where gender differences truly lie. The current research seeks to replicate previous findings regarding gender differences at the Big Five level, as well as to extend investigation into the intermediate sublevel of the two aspects within each domain.

Though no research has been done previously on gender differences at the aspect level of trait structure, we expect that the likely pattern of findings can be deduced from those reported for the Big Five and their facets. Because the aspects are more parsimonious and comprehensive than the facet models, however, they should provide a clearer and more systematic representation of gender differences in personality.

Our hypotheses were that women should score higher than men in both aspects of Neuroticism, Volatility, and Withdrawal, though the effect is likely to be stronger for Withdrawal, given the inclusion of anger within Volatility. Similarly, women should score higher than men in both aspects of Agreeableness, Compassion, and Politeness. Gender differences in the aspects of Conscientiousness, Industriousness, and Orderliness, may diverge, as research on facets suggests that women should score higher on Orderliness, but does not allow a clear prediction for Industriousness. The two aspects of Extraversion, Enthusiasm, and Assertiveness, should diverge because women should score higher than men in Enthusiasm (which combines sociability and positive emotionality), whereas men should score higher in Assertiveness. Gender differences should also be in opposite directions for the aspects of Openness/Intellect, Openness and Intellect. Women should score higher than men in Openness, whereas men should score higher than women in Intellect.

Use of the aspects has the additional advantage that one can easily examine the unique effects of one aspect while controlling for the other in each pair. In cases where gender differences on the two aspects diverge, this approach may reveal differences that are ordinarily suppressed by the shared variance of the two aspects within each Big Five domain. We implemented this approach through the use of residualized scores. Regressing one aspect on its complementary aspect and saving the residual produces a score that indicates unique variance in that aspect, without the variance it shares with its complement. For example, the residualized score for Compassion indicates differences in Compassion holding Politeness equal. If women are found, as predicted, to have higher Compassion residuals than men, that means that even if we take groups of men and women of equal Politeness, the women are nonetheless likely to be higher in Compassion on average.

Moderators

Due to the diversity of our sample, we performed secondary analyses to investigate potential moderators of gender differences. For example, previous research has shown that gender differences are larger and more pronounced within Western cultures than Eastern cultures (Costa et al., 2001; Schmitt et al., 2008). Though our sample was collected mostly within North America, we were interested if similar patterns would emerge when considering gender differences among people of different ethnic backgrounds. We were able to test whether the pattern of gender differences was similar in participants of European versus Asian ethnic backgrounds.

Additionally, previous research has shown that gender differences in some traits (such as negative affect) may be larger in emerging adulthood than in later adulthood (Soto et al., 2011). Therefore, we investigated whether age moderated the gender difference in each trait. Finally, an increasing number of studies are using an online method to administer personality measures. Our sample included both laboratory and online methods of administration. Though previous research has not shown significant difference in personality between these two methods (Gosling et al., 2004), we investigated whether administration method moderated gender differences in our sample.

MATERIALS AND METHODS

PARTICIPANTS

Participants ($N = 2643$; 892 male, 1751 female) were drawn from a number of research projects, for which they received either monetary compensation or university course credit. Much of the data was collected in a large Canadian metropolitan area, either as an online survey or as a part of laboratory studies ($N = 1826$; 537 male, 1289 female). Some participants ($N = 481$; 200 male, 281 female) were members of the Eugene-Springfield community sample (ESCS). Lastly, 336 participants were recruited via Amazon's Mechanical Turk (MTurk; 155 male, 181 female) and completed the measures online. Participants ranged in age from 17 to 85 ($M = 27.2$, SD = 14.4). The majority of participants identified as White (39.9%) or Asian (27.5%), with 1% or less identifying as Native American, Hispanic, and Black. Twenty-five percent of participants identified as "other," and 5% did not specify ethnicity. The demographic data for a number of our samples allowed participants to choose from only the above five ethnicity classifications or specify their ethnicity as "other." Therefore, the classification of "Asian" contains individuals of both South-Asian and East-Asian ethnic backgrounds. Though South-Asian and East-Asian cultures are markedly different in many ways, both are more collectivist than Western cultures (Suh et al., 1998) and therefore provide an interesting contrast to the White/European ethnic background.

PERSONALITY MEASURES

The Big Five aspect scales (BFAS) were designed specifically to assess the 10 aspects of the Big Five identified by DeYoung et al. (2007). Items were selected from the IPIP based on their correlation with the aspect factor scores and maintaining balanced keying. Items were chosen that differentiated the factor in question from all nine other aspect factors, by selecting items only if they were correlated with the aspect factor in question with a correlation at least 0.10 greater than the correlation with any other factor. This procedure has the consequence that the same items remain the best markers of each factor even when scores are residualized. Thus, the residualized scores retain the meaning of the construct in question. Ten items are used to assess each of the 10 aspects. Participants rate their agreement with how well each statement describes them using a five-point scale ranging from *strongly disagree* to *strongly agree*. Scores for each aspect are computed by taking the mean of the corresponding items. Scores for each domain are computed by taking the mean of the two aspect scores. The scales are all highly reliable (all $\alpha > 0.73$) and have good test–retest reliability, all $r > 0.72$ (DeYoung et al., 2007). Internal consistencies for the present data are shown in **Table 1**.

RESULTS

Table 2 summarizes the mean scores for men and women on each of the 10 aspects and the five domains. Because the two aspects within each domain are correlated (correlations range from $r = 0.39–0.62$), analyses on the aspects were performed on both raw scores and residualized scores. Residualized scores were created by regressing one aspect within a domain on the other, and saving the residuals, thus creating an index of the variance of each aspect not associated with its complement in the same domain. For example,

the residualized scores for Enthusiasm are the residuals resulting from the regression of Enthusiasm on Assertiveness, effectively partialing out the general Extraversion variance shared between Enthusiasm and Assertiveness. **Table 3** presents the intercorrelations between aspects, analyzed separately for men and women.

Gender differences were analyzed using independent samples t-tests. Effect sizes are summarized in **Table 2**. Results were consistent with previous analyses, with significant effects found at the level of the Big Five domains of Neuroticism, Agreeableness, and Extraversion, but not Conscientiousness or Openness. The largest effect sizes were found for Neuroticism and Agreeableness. Unsurprisingly, Neuroticism and Agreeableness are the domains for which gender differences were significant and in the same direction for both underlying aspects.

Using the raw scores, gender differences were found in all of the aspects with the exception of Industriousness. Women scored higher than men on Enthusiasm, Compassion, Politeness, Orderliness, Volatility, Withdrawal, and Openness. Men scored higher than women on Assertiveness and Intellect. This indicates that the two aspects of Extraversion (Enthusiasm and Assertiveness) and the two aspects of Openness/Intellect display gender differences in opposite directions. Such divergence in gender differences at the aspect level helps to elucidate the small effect of the gender difference in overall Extraversion and the lack of a significant gender difference in Openness/Intellect. As in previous research, effect sizes were small to moderate (range: 0.06–0.48 in absolute value).

Results for the residualized scores differed from those on the raw scores in two ways. First, the gender difference in Industriousness was now significant, with men scoring higher than women. Since this is a residualized score, it indicates a gender difference in Industriousness among people with equal levels of Orderliness. Second, there was not a significant gender difference in residualized

Table 1 | Alpha reliabilities for Big Five domains and aspect scales by sample.

	Canadian	ESCS	MTurk
Sample size	1827	481	336
Enthusiasm	0.82	0.81	0.79
Assertiveness	0.85	0.85	0.83
Compassion	0.83	0.84	0.79
Politeness	0.73	0.75	0.76
Industriousness	0.81	0.81	0.81
Orderliness	0.75	0.80	0.81
Volatility	0.88	0.85	0.88
Withdrawal	0.81	0.84	0.81
Intellect	0.79	0.84	0.68
Openness	0.74	0.78	0.71
Extraversion	0.88	0.85	0.86
Agreeableness	0.83	0.84	0.85
Conscientiousness	0.82	0.84	0.86
Neuroticism	0.89	0.89	0.90
Openness/Intellect	0.81	0.85	0.78
Aspect average	0.83	0.82	0.79
Domain average	0.85	0.85	0.85

Table 2 | Mean and SD for Big Five domains and raw and residualized aspect scores.

| | Raw scores | | | | | Residualized scores | | | | |
| | Males | | Females | | | Males | | Females | | |
	Mean	SD	Mean	SD	*d*	Mean	SD	Mean	SD	*d*
Enthusiasm	3.40	0.66	3.56	0.68	**0.23**	−0.13	0.59	0.06	0.59	**0.32**
Assertiveness	3.34	0.64	3.28	0.68	**−0.09**	0.09	0.57	−0.05	0.58	**−0.24**
Compassion	3.78	0.60	4.04	0.56	**0.45**	−0.11	0.56	0.05	0.49	**0.31**
Politeness	3.52	0.61	3.74	0.61	**0.36**	−0.06	0.57	0.04	0.53	**0.18**
Industriousness	3.25	0.68	3.21	0.73	−0.06	0.07	0.62	−0.03	0.66	**−0.15**
Orderliness	3.38	0.61	3.49	0.63	**0.18**	−0.08	0.56	0.05	0.57	**0.22**
Volatility	2.63	0.75	2.86	0.77	**0.30**	−0.03	0.60	0.01	0.61	0.06
Withdrawal	2.73	0.69	3.02	0.71	**0.40**	−0.10	0.55	0.05	0.56	**0.27**
Intellect	3.62	0.61	3.48	0.63	**−0.22**	0.14	0.57	−0.07	0.58	**−0.36**
Openness	3.57	0.61	3.74	0.60	**0.27**	−0.14	0.56	0.07	0.54	**0.39**
Extraversion	3.37	0.55	3.42	0.59	**0.08**					
Agreeableness	3.65	0.50	3.89	0.50	**0.48**					
Conscientiousness	3.32	0.54	3.35	0.58	0.06					
Neuroticism	2.68	0.65	2.94	0.67	**0.39**					
Openness/Intellect	3.60	0.51	3.61	0.52	0.02					

Bolded d values indicate statistically significant effect sizes.

Table 3 | Correlations among aspects (raw scores).

	Enth.	Assert.	Comp.	Polit.	Indust.	Order.	Volat.	Withd.	Intel.	Open.
Enthusiasm	–	0.45	0.46	0.10	0.21	−0.00	−0.18	−0.35	0.17	0.15
Assertiveness	0.51	–	0.22	−0.20	0.34	0.06	−0.09	−0.40	0.47	0.20
Compassion	0.45	0.25	–	0.38	0.14	0.03	−0.17	−0.12	0.24	0.43
Politeness	0.12	−0.17	0.47	–	0.29	0.23	−0.35	−0.20	−0.01	0.07
Industriousness	0.22	0.36	0.21	0.28	–	0.41	−0.24	−0.51	0.33	−0.02
Orderliness	0.04	0.08	0.14	0.21	0.44	–	0.03	−0.03	−0.02	−0.07
Volatility	−0.15	−0.07	−0.15	−0.40	−0.36	−0.02	–	0.61	−0.23	0.00
Withdrawal	−0.37	−0.42	−0.14	−0.17	−0.54	−0.03	0.62	–	−0.35	0.03
Intellect	0.22	0.48	0.31	0.10	0.42	0.04	−0.21	−0.35	–	0.38
Openness	0.13	0.17	0.39	0.17	0.01	−0.03	0.01	0.04	0.42	–

Correlations above the diagonal are for males; females are below the diagonal.

scores of Volatility. This indicates no difference between the average scores of men and women in Volatility when they are at equal levels of Withdrawal.
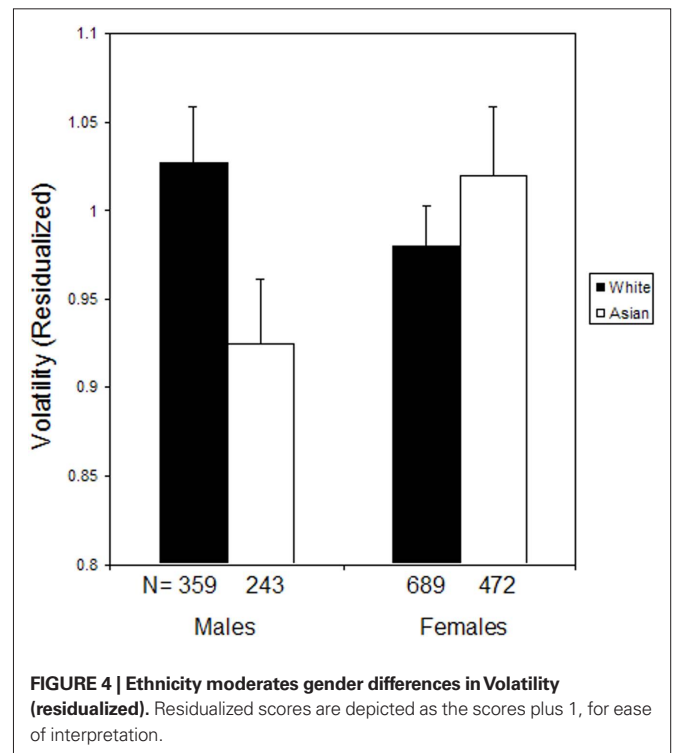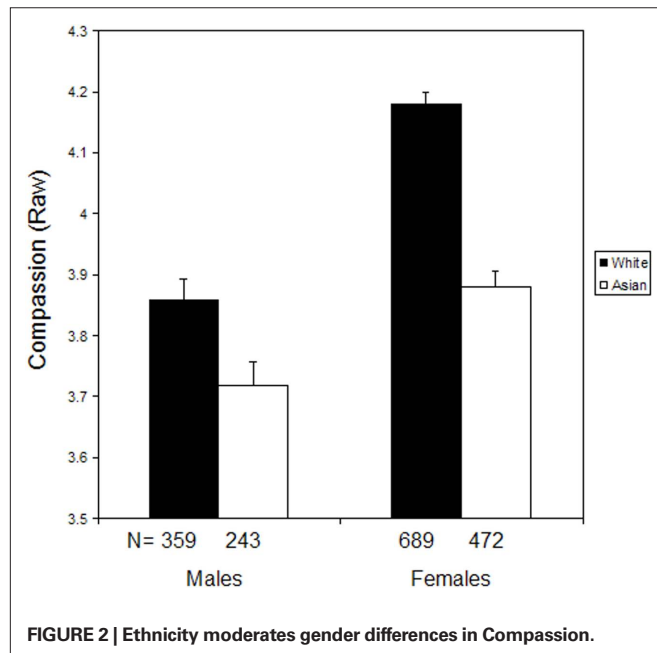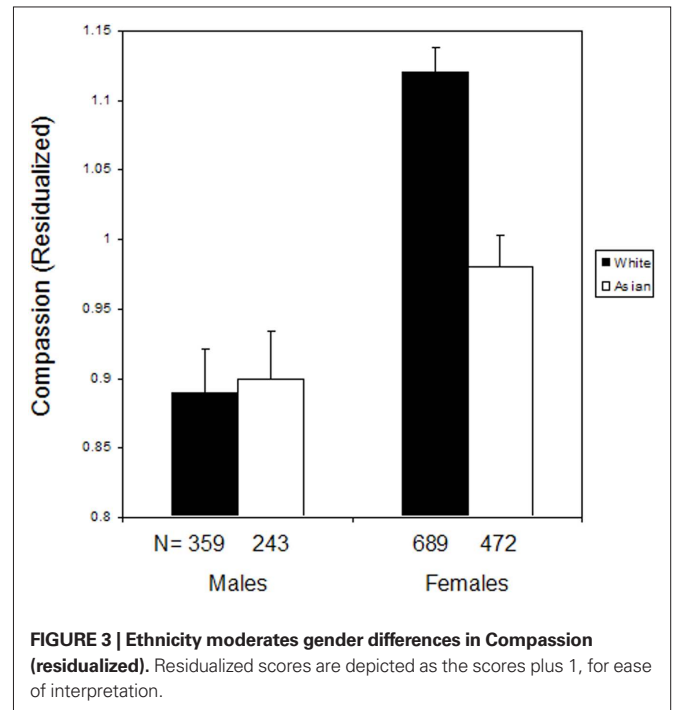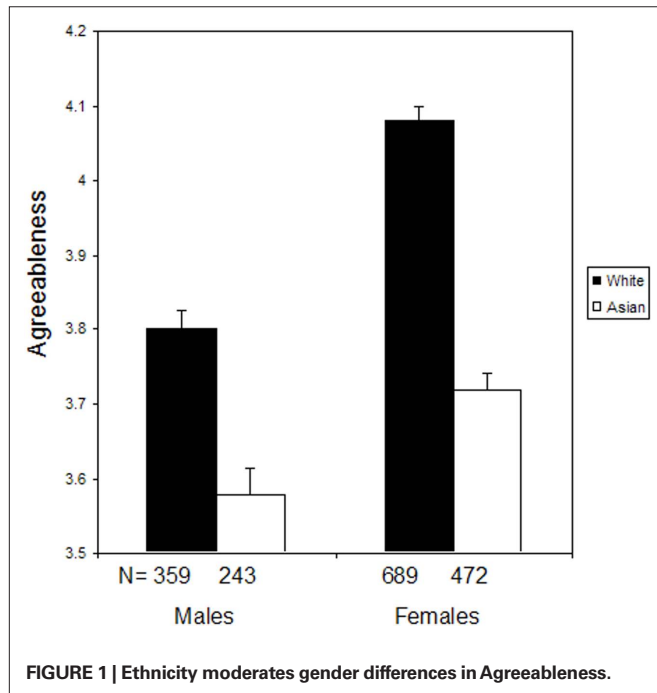
MODERATORS

We performed regression analyses to see if ethnicity, age, and survey method moderated the gender differences we had found. Previous research suggests that gender differences are robust across cultures (Costa et al., 2001; McCrae et al., 2005), and may differ with age for some traits (Soto et al., 2011). Because the majority of the participants in our sample were White and Asian, we were able to make comparisons only for these two groups.

Ethnicity (coded as White or Asian) significantly moderated gender differences in the Big Five domain of Agreeableness, $F(1, 1759) = 5.42$, $p = 0.02$, with a more pronounced gender difference found among Whites than among Asians (see **Figure 1**). Similar patterns were found for ethnicity moderating the gender difference in Compassion in both the raw scores, $F(1, 1759) = 7.97$, $p < 0.001$, (see **Figure 2**) and the residualized scores, $F(1, 1759) = 6.64$, $p = 0.01$, (see **Figure 3**). When Politeness was partialed out of the Compassion scores, there was no difference between White men and Asian men and a significant difference between White women and Asian women, such that the gender difference was more pronounced for Whites than for Asians.

Ethnicity also moderated gender differences in the residualized scores for Volatility, $F(1, 1759) = 5.09$, $p = 0.02$, though this pattern was somewhat different. The gender difference was significant for Asian participants such that women scored higher than men. However, for White participants, men scored higher than women (see **Figure 4**). No other moderations by ethnicity were observed.

FIGURE 1 | Ethnicity moderates gender differences in Agreeableness.

FIGURE 2 | Ethnicity moderates gender differences in Compassion.

FIGURE 3 | Ethnicity moderates gender differences in Compassion (residualized). Residualized scores are depicted as the scores plus 1, for ease of interpretation.

FIGURE 4 | Ethnicity moderates gender differences in Volatility (residualized). Residualized scores are depicted as the scores plus 1, for ease of interpretation.

Age significantly moderated gender differences in the Big Five domains of Agreeableness, $F(1, 2576) = 4.88$, $p = 0.03$, Neuroticism, $F(1, 2576) = 11.35$, $p < 0.001$, and Openness, $F(1, 2576) = 4.26$, $p = 0.04$. The gender difference in Agreeableness was larger for older ages, and the gender difference in Neuroticism was larger for younger ages. In addition, the gender difference seemed to reverse for Neuroticism, such that men had higher scores than women in older ages. For Openness/Intellect, the gender difference was non-existent at younger ages, and larger favoring women at older ages. These patterns were driven by spe-

cific aspects, as evidenced by age's moderating the gender difference in Compassion, $F(1, 2576) = 7.64$, $p = 0.01$ (see **Figure 5**), Volatility, $F(1, 2576) = 19.79$, $p < 0.001$ (see **Figure 6**), and Intellect, $F(1, 2576) = 3.96$, $p = 0.05$ (see **Figure 7**). The pattern for Intellect is such that there is a larger gender difference for younger ages than for older, favoring men. At older ages, the gender difference is non-existent or slightly favors women.

Age also moderated gender differences in the residualized scores for Compassion, $F(1, 2576) = 6.76$, $p = 0.01$, Orderliness, $F(1,2576) = 5.02$, $p = 0.03$ (see **Figure 8**), and Volatility $F(1, 2576) = 20.21$, $p < 0.001$. The pattern for Compassion was similar to that found for the raw scores, such that the gender difference in residualized Compassion increased with age. The gender difference in residualized Orderliness was small and favored women at younger ages, yet it decreased to non-existence and almost reversed to favor men in older ages. Finally, the pattern for Volatility was similar to that found for the raw scores, such that the gender difference favored women at younger ages and men at older.

The format of the administration of the surveys moderated gender differences in Compassion on both the raw, $F(1, 2115) = 6.14$, $p = 0.01$ (see **Figure 9**), and residualized, $F(1, 2115) = 6.63$, $p = 0.01$, metrics. Womens' Compassion scores did not differ between the online and laboratory format. Men who completed the survey in the lab had higher average Compassion scores than did men who completed the survey online. The results for the residualized Compassion scores were nearly identical to those for the raw scores, depicted in **Figure 9**.

DISCUSSION

Gender differences were more pervasive at the aspect level of trait organization immediately below the Big Five than for the Big Five themselves. At the level of the Big Five our findings were similar to the typical pattern: gender differences were found only for Neuroticism, Agreeableness, and Extraversion. However, gender differences were found for every one of the 10 aspects, considering analyses of both raw and residualized scores. Even when examining raw scores alone, differences were significant in 9 out of 10 traits. Clearly, analysis of the aspects reveals the extent of gender differences across the whole spectrum of traits encompassed by the hierarchical Big Five model.
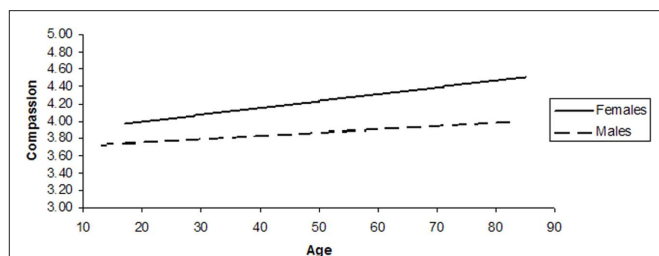

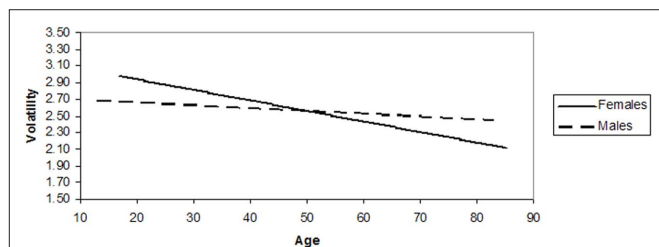FIGURE 5 | Age moderates gender differences in Compassion.

FIGURE 6 | Age moderates gender differences in Volatility.

NEUROTICISM

Consistent with previous findings, women scored higher than men in Neuroticism and in both of its aspects, Withdrawal and Volatility, when measured in terms of raw scores. This difference replicates previous findings for Neuroticism (e.g., Costa et al., 2001). Because Withdrawal and Volatility are correlated ($r = 0.62$ in our sample as a whole) but distinct traits, it is important to consider the unique variance that each does not share with the other. We therefore additionally examined gender
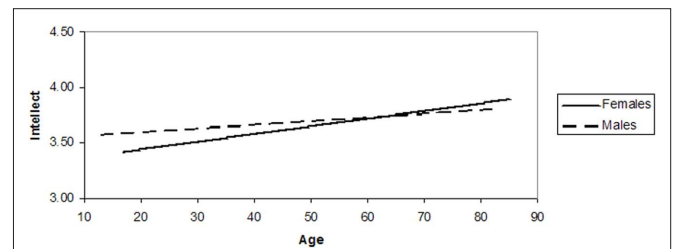

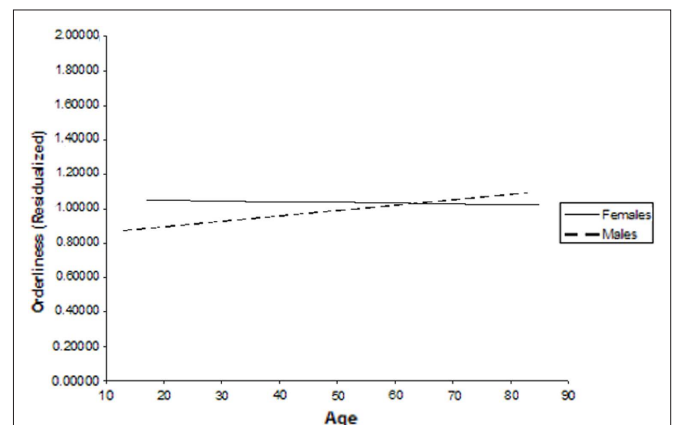FIGURE 7 | Age moderates gender differences in Intellect.

FIGURE 8 | Age moderates gender differences in Orderliness (residualized). Residualized scores are depicted as the scores plus 1, for ease of interpretation.
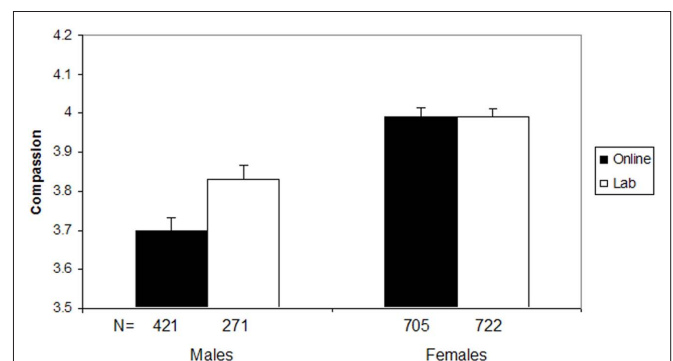

FIGURE 9 | Administration method moderates gender differences in Compassion.

differences in Withdrawal with Volatility partialed out, and *vice versa*. The gender difference remained for Withdrawal but was eliminated for Volatility. This contrast for the unique variance in these two aspects is not surprising if one considers the more specific content of each. At the facet level of Neuroticism, women have been found to show higher levels of anxiety, depression, self-consciousness, and vulnerability than men (Costa et al., 2001). All of these facets load primarily on Withdrawal rather than Volatility (DeYoung et al., 2007). This pattern is consistent with the fact that clinical diagnoses of depression and anxiety are considerably more common in women than men (Weissman et al., 1996).

In contrast, the lack of a significant gender difference in Volatility, when controlling for Withdrawal, is most likely due to the fact that an important component of Volatility is the tendency to be irritable and easily angered. Men have sometimes been found to score higher than women on traits such as Anger or Hostility (Scherwitz et al., 1991).

The gender difference in Neuroticism was moderated by age, such that the gender difference decreased with age. Neuroticism increases during emerging adulthood among females, but not males (Soto et al., 2011), which may explain this pattern of results. The gender difference in Volatility was moderated by both age and ethnicity. For age, the pattern was the same as for the overall domain of Neuroticism, and was seen for both Volatility on the raw scale metric and when controlling for Withdrawal. The moderating effect ethnicity had on residualized scores for Volatility was different, showing gender differences in opposite directions among White and Asian participants. Men scored higher in Volatility than women among White participants, whereas women scored higher among Asian participants. Given the fact that Volatility partly reflects traits related to irritability and anger, this difference may be due to cultural differences in social norms related to the expression of anger (Matsumoto and Fontaine, 2008).

AGREEABLENESS

Replicating previous findings, there was a significant gender difference in Agreeableness such that women tend to score higher than men, and this pattern was the same for the aspects, Compassion and Politeness, when measured in terms of raw scores or residualized scores. Compassion most clearly represents a tendency to invest in others emotionally and affiliate on an emotional level, encompassing traits such as warmth and empathy. Politeness describes the tendency to show respect to others and refrain from taking advantage of them, and is related to traits such as cooperation and compliance. Our findings that women score higher than men on both aspects are consistent with previous research showing women are more trusting and compliant than men (Costa et al., 2001).

Gender differences in Agreeableness may be related to gender differences in self-construal. Men tend to have an independent self-construal, or a sense of self that is separate from cognitive representations of others. Women have a more interdependent self-construal, in which their sense of self includes others (Markus and Kitayama, 1991). This gender difference is associated with motivational and behavioral differences, such as women having more interconnected and affiliative social groups (Cross and

Madson, 1997). Women, therefore, may be more motivated than men to maintain social and emotional bonds by enacting more agreeable traits.

Age moderated gender differences in Agreeableness and Compassion on both the raw and residualized metric, such that gender differences were larger among older individuals.

Ethnicity moderated the gender differences in Agreeableness and its Compassion aspect, such that differences were larger among White participants than among Asian participants. This finding is consistent with previous research, which shows larger gender differences among more western and industrialized cultures (Costa et al., 2001; McCrae et al., 2005; Schmitt et al., 2008). Asian participants in general rated themselves as less agreeable than white participants, which may indicate a reference group effect. A reference group effect would occur if participants are comparing themselves to their own culture and there is a difference in Agreeableness between cultures. For example, someone who is more agreeable than the norm for Whites may be less agreeable than the norm for Asians. Asian participants could be comparing themselves to an allocentric cultural norm, in which consideration for others is central and therefore higher Agreeableness is normative (Triandis, 2001).

Method of administration of the measure also moderated the gender difference in Compassion on both the raw and residualized metric. The gender difference was larger in online administration than for laboratory administration. This is because men scored higher in Compassion on average when they completed the measure in the laboratory than when they completed it online. This may be due to social desirability effects' causing men to report being higher in Compassion when they are in the laboratory, and not when they are completing the measure online.

CONSCIENTIOUSNESS

Consistent with previous research, we did not find a significant gender difference in Conscientiousness at the level of the Big Five domain. When measuring the aspects of Industriousness and Orderliness in terms of raw scores, however, we found a significant gender difference for Orderliness, such that women score higher than men. The Orderliness aspect reflects traits related to maintaining order and organization, including perfectionism (DeYoung et al., 2007). Given the positive correlation between perfectionism and components of Neuroticism such as anxiety and depression (Dunkley et al., 2006; Sherry et al., 2007), and the well-established gender differences in Neuroticism, one possibility is that Neuroticism accounts for the gender difference in Orderliness. However, when we regressed Orderliness (either raw or residualized) on Neuroticism and gender, gender remained a significant predictor, indicating that gender differences in Orderliness are not simply due to differences in Neuroticism.

Age moderated the gender difference in residualized Orderliness, such that the gender difference favoring women seen at younger ages decreased to non-existence and reversal at older ages. The age trend for women indicated a decline in Orderliness relative to Industriousness, whereas the trend for men indicated an increase.

Though no gender difference was found for Industriousness when using the raw scores, we found a gender difference in Industriousness when using the residualized score that removed

any variance shared with Orderliness. This gender difference was such that men scored higher than women in Industriousness. This difference in residualized scores but not raw scores can be interpreted as follows: if one examines a group of people with equal levels of Orderliness, the men in that group will on average score higher in Industriousness than the women.

EXTRAVERSION

We found a small but significant gender difference in overall Extraversion such that women score higher than men. However, the pattern was more complicated for the aspects, Enthusiasm and Assertiveness. Enthusiasm reflects sociability, gregariousness, and experiences of positive emotion. Our finding that women score higher than men in Enthusiasm was consistent with previous research showing similar patterns in Big Five facets of Gregariousness and Positive Emotions (Feingold, 1994; Costa et al., 2001). Assertiveness, on the other hand, reflects traits related to agency and dominance. Consistent with previous research showing a gender difference favoring men for facets such as Assertiveness and Excitement Seeking (Feingold, 1994; Costa et al., 2001), we found that men score higher than women in Assertiveness. This pattern of gender differences in opposite directions at the aspect level was also found for scores on the residualized metric.

OPENNESS/INTELLECT

Consistent with previous research, we did not find a significant difference in Openness/Intellect at the level of the Big Five domain. However, we found significant gender differences in both aspects of the Big Five domain (Intellect and Openness). On both the raw and residualized scores women scored higher than men in Openness. In contrast, on both types of score, men scored higher than women in Intellect. This pattern is consistent with previous reports of gender differences at the facet level, where women score higher than men on facets marking Openness (such as Esthetics and Feelings), but men score higher than women on the Ideas facet, which is a marker of Intellect (Feingold, 1994; Costa et al., 2001).

Although Intellect includes perceptions of cognitive ability and is more strongly associated with IQ scores than Openness (DeYoung et al., submitted), the fact that men score higher than women in Intellect should not be taken as indicative of greater intelligence for men than women. Gender differences in general intelligence are negligible, although men are typically found to show more variance in scores than women (Deary et al., 2007; van der Sluis et al., 2008). However, our findings are consistent with the finding that men show higher self-estimates of intelligence than women, across cultures (von Stumm et al., 2009). This pattern has been described as one of male hubris and female humility in relation to intelligence. The gender difference in Intellect probably reflects these biases related to confidence in intellectual abilities.

Age moderated the gender difference in Intellect such that the gender difference was smaller at higher ages. The pattern suggests that the difference in Intellect between older and younger women is larger than that between older and younger men. Since gender differences in intelligence are negligible across the lifespan, this pattern most likely indicates that women gain in perceptions of their own intelligence, perhaps reflecting increases in self-esteem or self-confidence (Orth et al., 2010).
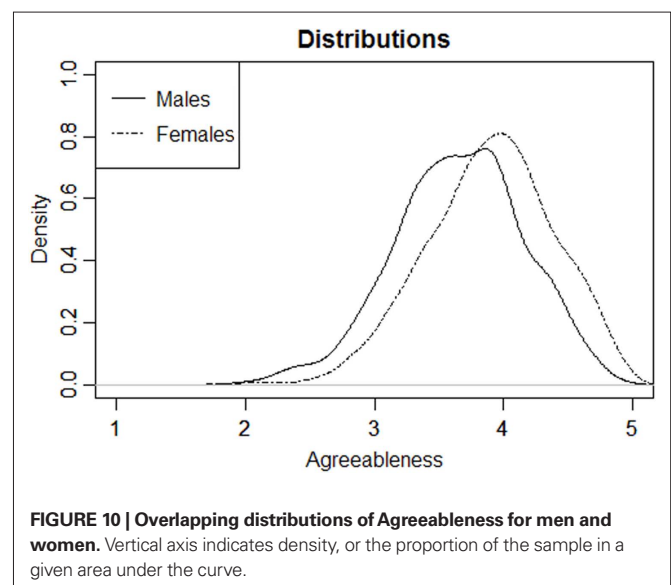
LIMITATIONS AND FUTURE DIRECTIONS

Our investigation was limited to one measure of personality, the BFAS. Although the Big Five organization of personality that it employs is reasonably comprehensive, there are traits that may not have good representation among the items in the BFAS, such as adult attachment style (Hazan and Shaver, 1987). It would be worthwhile for future research to investigate gender differences in these additional traits as well as how they relate to gender differences in the Big Five.

Further, the personality scores used in our investigation were obtained via self-report. Our findings could therefore indicate gender differences in how men and women perceive and report on themselves, which do not necessarily reflect how they are perceived by others or their actual behavioral tendencies. Future research should explore gender differences in peer-reports of these personality traits. This might be especially interesting when the perceiver is of the opposite gender from the target. Additionally, behavioral or implicit measures of personality could be used to investigate whether the same pattern of gender differences exist when one moves beyond measuring personality through questionnaires.

Previous research has investigated gender differences among many different ethnicities, cultures, and types of societies (McCrae et al., 2005; Schmitt et al., 2008). Such an investigation is beyond the scope of the current research. The current sample was mainly North American, and sample sizes within each ethnic group were limited, such that we were only able to perform analyses comparing White participants to Asian participants. It would be beneficial for future research to investigate gender differences in personality at the aspect level in additional ethnicities and cultures.

Similarly, our research indicated that age moderated gender differences in a number of traits. Our sample was cross-sectional rather than longitudinal, hence our results may not accurately reflect the trajectories of personality change in men and women over time. Taken along with previous findings on age trends in personality (e.g., Roberts et al., 2006; Soto et al., 2011), our results suggest the utility of further investigation of how gender differences in personality may differ with increasing age.

FIGURE 10 | Overlapping distributions of Agreeableness for men and women. Vertical axis indicates density, or the proportion of the sample in a given area under the curve.

Finally, though this and other studies have shown the existence of gender differences in personality, the question remains as to why these differences exist. Although the general consistency of gender differences across cultures may suggest evolutionary reasons for the existence of gender differences in personality traits, cross-cultural variation in gender differences for some trait may suggest that culture of origin or social roles and norms influence gender differences. Exactly how culture impacts personality is a complex question, worthy of future study.

CONCLUSION

By examining personality at the level of the 10 aspects of the Big Five, we demonstrated that gender differences in personality traits are even more pervasive than has typically been reported. In every one of the 10 traits assessed, significant gender differences were evident. For some Big Five domains, the aspect level traits showed gender differences in opposite directions, which helps to explain why gender differences are not typically evident for the Big Five domains of Conscientiousness and Openness/Intellect, and why the gender difference for Extraversion is typically very small.

Clearly the average personalities of men and women are systematically different. Does this mean, however, that Bill Cosby's metaphor, that men and women are from "different species," is apt? We would caution against adopting such a dramatic interpretation of the pervasive gender differences in personality that we report in this study. All of the mean differences we found (and all of the differences that have been found in the past – e.g., Feingold, 1994; Costa et al., 2001) are small to moderate. This means that the distributions of traits for men and women are largely overlapping. To illustrate this fact, in **Figure 10** we present the male and female distributions from our sample for the trait which showed the largest gender difference, Agreeableness. One can see that both men and women can be found across a similar range of Agreeableness scores, such that, despite the fact that women score higher than men on average, there are many men who are more agreeable than many women, and many women who are less agreeable than many men. Given that Agreeableness showed the largest gender difference in our study, all other traits for which we reported significant gender differences would show even greater overlap in men's and women's distributions. Although the mean differences in personality between genders may be important in shaping human experience and human culture, they are probably not so large as to preclude effective communication between men and women. Unlike Bill Cosby, we are optimistic that any difficulties in communication between men and women are due primarily to cultural norms that are amenable to change, rather than to differences in basic personality traits, which are much more difficult to change.

REFERENCES

Buss, D. M. (2008). *Evolutionary Psychology: The New Science of the Mind*, 3rd Edn. Boston: Allyn & Bacon.

Costa, P. T. Jr., and 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. Jr., Terracciano, A., and McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *J. Pers. Soc. Psychol.* 81, 322–331.

Cross, S. E., and Madson, L. (1997). Models of the self: self-construals and gender. *Psychol. Bull.* 122, 5–37.

Deary, I. J., Irwing, P., Der, G., and Bates, T. C. (2007). Brother-sister differences in the g factor in intelligence: analysis of full, opposite-sex siblings from the NLSY1979. *Intelligence* 35, 451–456.

Depue, R. A., and Collins, P. F. (1999). Neurobiology of the structure of personality: dopamine, facilitation of incentive motivation, and extraversion. *Behav. Brain Sci.* 22, 491–569.

DeYoung, C.G., and Gray, J.R. (2009). "Personality neuroscience: explaining individual differences in affect, behavior, and cognition," in *The Cambridge Handbook of Personality Psychology*, eds P.J. Corr and G. Matthews (New York: Cambridge University Press), 323–346.

DeYoung, C. G., Quilty, L. C., and Peterson, J. B. (2007). Between facets and domain: 10 aspects of the Big Five. *J. Pers. Soc. Psychol.* 93, 880–896.

Digman, J. M. (1990). Personality structure: emergence of the five-factor model. *Annu. Rev. Psychol.* 41, 417–440.

Dunkley, D. M., Sanislow, C. A., Grilo, C. M., and McGlashan, T. H. (2006). Perfectionism and depressive symptoms 3 years later: negative social interactions, avoidant coping, and perceived social support as mediators. *Compr. Psychiatry* 46, 106–115.

Eagly, A. H., and Wood, W. (2005). Universal sex differences across patriarchal cultures ≠ evolved psychological dispositions. *Behav. Brain Sci.* 28, 281–283.

Feingold, A. (1994). Gender differences in personality: a meta-analysis. *Psychol. Bull.* 116, 429–456.

Fleeson, W., and Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: fifteen experience-sampling studies and a meta-analysis. *J. Pers. Soc. Psychol.* 97, 1097–1114.

Gosling, S. D., Vazire, S., Srivastava, S., and John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *Am. Psychol.* 59, 93–104.

Hazan, C., and Shaver, P. (1987). Romantic love conceptualized as an attachment process. *J. Pers. Soc. Psychol.* 52, 511–524.

Helgeson, V. S., and Fritz, H. L. (1999). Unmitigated agency and unmitigated communion: distinctions from agency and communion. *J. Res. Pers.* 33, 31–158.

Hyde, J. S. (2005). The gender similarities hypothesis. *Am. Psychol.* 60, 581–592.

Jang, K. L., Livesley, W. J., Angleitner, A., Reimann, R., and Vernon, P. A. (2002). Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. *Pers. Individ. Dif.* 33, 83–101.

John, O. P., Naumann, L. P., and Soto, C. J. (2008). "Paradigm shift to the integrative Big Five trait taxonomy: history, measurement, and conceptual issues," in *Handbook of Personality Psychology: Theory and Research*, 3rd Edn, eds O. P. John, R. W. Robins, and L. A. Pervin (New York: Guilford Press), 114–158.

Kling, K. C., Hyde, J. S., Showers, C. J., and Buswell, B. N. (1999). Gender differences in self-esteem: a meta-analysis. *Psychol. Bull.* 125, 470–500.

Lippa, R. A. (2010). Gender differences in personality and interests: when, where, and why? Social and personality psychology compass. 4, 1098–1110.

Markus, H., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253.

Matsumoto, D., and Fontaine, J. (2008). Mapping expressive differences around the world: the relationship between emotional display rules and individualism versus collectivism. *J. Cross Cult. Psychol.* 39, 55–74.

McCrae, R. R., and Costa, P. T. Jr. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *J. Pers. Soc. Psychol.* 56, 586–595.

McCrae, R. R., Terracciano, A., and 79 members of the Personality Profiles of Cultures Project. (2005). Personality profiles of cultures: aggregate personality traits. *J. Pers. Soc. Psychol.* 89, 407–425.

Neisser, U., Boodoo, G., Bouchard, T., Boykin, A., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., and Urbina, S. (1996). Intelligence: knowns and unknowns. *Am. Psychol.* 51, 77–101.

Orth, U., Trzesniewski, K. H., and Robins, R. W. (2010). Self-esteem development from young adulthood to old age: a cohort-sequential longitudinal study. *J. Pers. Soc. Psychol.* 98, 645–658.

Roberts, B. W., Walton, K. E., and Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta analysis of longitudinal studies. *Psychol. Bull.* 132, 1–25.

Scherwitz, L., Perkins, L., Chesney, M., and Hughes, G. (1991). Cook-Medley Hostility scale and subsets: relationship to demographic and psychosocial characteristics in young adults in the CARDIA study. *Psychosom. Med.* 53, 36–49.

Schmitt, D. P., Realo, A., Voracek, M., and Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *J. Pers. Soc. Psychol.* 94, 168–182.

Sherry, S. B., Hewitt, P. L., Flett, G. L., Lee-Baggely, D. L., and Hall, P. A. (2007). Trait perfectionism and perfectionistic self-presentation in personality pathology. *Pers. Individ. Dif.* 42, 477–490.

Soto, C. J., and John, O. P. (2009). Using the California psychological inventory to assess the Big Five personality domains: a hierarchical approach. *J. Res. Pers.* 43, 25–38.

Soto, C. J., John, O. P., Gosling, S. D., and Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *J. Pers. Soc. Psychol.* 100, 330–348.

Suh, E., Diener, E., Oishi, S., and Triandis, H. C. (1998). The shifting basis of life satisfaction judgments across cultures: emotions versus norms. *J. Pers. Soc. Psychol.* 74, 482–493.

Triandis, H. C. (2001). Individualism-collectivism and personality. *J. Pers.* 69, 907–924.

Trivers, R. L. (1972). "Parental investment and sexual selection," in *Sexual Selection and the Descent of Man*, ed. B. Campbell (Chicago: Aldine), 136–179.

van der Sluis, S., Derom, C., Thiery, E., Bartels, M., Polderman, T. J. C., Verhulst, F. C., Jacobs, N., van Gestel, S., de Geus, E. J. C., Dolan, C. V., Boomsma, D. I., and Posthuma, D. (2008). Sex differences on the WISC-R in Belgium and the Netherlands. *Intelligence* 36, 48–67.

von Stumm, S., Chamorro-Premuzic, T., and Furnham, A. (2009). Decomposing self-estimates of intelligence: structure and sex differences across 12 nations. *Br. J. Psychol.* 100, 429–442.

Weissman, M. M., Bland, R. C., Canino, G. J., Faravelli, C., Greenwald, S., Hwu, H. G., Joyce, P. R., Karam, E. G., Lee, C. K., Lellouch, J., Lépine, J. P., Newman, S. C., Rubio-Stipec, M., Wells, J. E., Wickramaratne, P. J., Wittchen, H., and Yeh, E. K. (1996). Cross-national epidemiology of major depression and bipolar disorder. *J. Am. Med. Assoc.* 276, 293–299.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: the interpersonal domain. *J. Pers. Soc. Psychol.* 37, 395–412.

Wood, W., and Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: implications for the origins of sex differences. *Psychol. Bull.* 128, 699–727.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.