Sex differences in 30 facets of the five factor model of personality in the large public (N = 320,128)

Petri J. Kajonius¹,b,c,* , John Johnson⁴

¹ Department of Social and Behavioral Studies, University West, Sweden
² Department of Psychology, University of Gothenburg, Sweden
³ Department of Cognitive Neuroscience, University of Skövde, Sweden
⁴ Department of Psychology, Pennsylvania State University, USA

ARTICLE INFO

Keywords:
Personality traits
Sex
Measurement
Five factor model

ABSTRACT

The present study reports on the scope and size of sex differences in 30 personality facet traits, using one of the largest US samples to date (N = 320,128). The study was one of the first to utilize the open access version of the Five-Factor Model of personality (IPIP-NEO-120) in the large public. Overall, across age-groups 19–69 years old, women scored notably higher than men in Agreeableness (d = 0.58) and Neuroticism (d = 0.40). Specifically, women scored d > 0.50 in facet traits Anxiety, Vulnerability, Openness to Emotions, Altruism, and Sympathy, while men only scored slightly higher (d > 0.20) than women in facet traits Excitement-seeking and Openness to Intellect. Sex gaps in the five trait domains were fairly constant across all age-groups, with the exception for age-group 19–29 years old. The discussion centers on how to interpret effects sizes in sex differences in personality traits, and tentative consequences.

1. Psychological sex differences

One of the most informative meta-analyses on psychological differences between sexes encompasses an impressive 100 meta-analyses, and concludes that the overall difference between sexes is small (Cohen’s d = 0.21) (Zell, Teeter, & Zell, 2015). An effect size of 0.2 standard deviations (d) would translate into approximately 58% of one sex having a higher score than the mean of the other (Cohen, 1992; See http://rpsychologist.com/d3/cohen/ for calculations). Nevertheless, large apparent dissimilarities were also reported, such as mental rotation abilities (d = 0.57), and attitudes towards mate-beauty (d = 0.53). In general, sexual and physical aggression tends to be male dominated traits (e.g. Lippa, 2009). Along the same line, traits such as risk-taking and power motivation tend to be male-dominated (d > 0.40) (Byrnes, Miller, & Schafer, 1999), while affiliation motivation (d = 0.45) tend to be more female-dominated (Drescher & Schultheiss, 2016). In general, dispositions for care and benevolence are found to be higher in females (d = 0.28) (Jaffee & Hyde, 2000), and are considered female traits across 127 cultures (Schwartz & Rubel, 2005). Sex differences concerning personality traits have since long been established as be small to moderate, but consequential (Feingold, 1994).
Personality and Individual Differences 129 (2018) 126–130

P.J. Kajonius, J. Johnson

2. Sex differences in personality traits

Personality traits are defined as consistencies in affect, thinking, and behavior that tend to be stable across situations and time-spans, developing predictably throughout life (Briley & Tucker-drob, 2014). The most conceptualized model is today the Five Factor Model (FFM) (McCrae, 2010). This model organizes personality into five trait factor domains (N = Neuroticism, E = Extraversion, O = Openness, A = Agreeableness, and C = Conscientiousness), which in turn are composed by a number of specific trait facets. Facets help increase the precision and scope of personality, thus enabling more accurate predictions (Ziegler & Bäckström, 2016). For instance, knowing that someone scores high on trait factor Openness may not give us meaningful information on how this person would fare in specific circumstances. Instead, a specific facet such as Imagination may be more informative for instance in career choice, or facet Intellect in educational settings. Furthermore, facets may yield more information than trait domains when comparing groups. For instance, the trait factor Extraversion usually does not differ much between sexes, not even in adolescence (Borghuis et al., 2017); however, a facet underlying this such as Excitement-seeking is usually notably higher with males. See a full list of differences between sexes across cultures (Costa Jr et al., 2010). Not too many other studies have utilized facets of personality in very large samples across 22 nations (Kajonius & Mac Giolla, 2017). Not only does the FFM trait facets, seems to replicate cross-culturally, as reported in large samples across 22 nations (Kajonius & Mac Giolla, 2017). Not too many other studies have utilized facets of personality in very large samples. The consensus today is that sex differences in FFM are often small but significant, and that males tend to vary more in personality traits than females (Borkenau, McCrae, & Terracciano, 2013). Also, the literature also portrays some agreement in that the more individualistic and equal society is, the larger the demonstrated sex differences (Costa Jr et al., 2001; Schmitt, Realo, Voracek, & Allik, 2008). In addition, this effect may be growing with increased individual freedom (cf. Skirbekk & Blekesaune, 2014). Sex differences also develop predictably across the life-span. Already in early adolescence both Agreeableness and Neuroticism traits are found to be higher in girls, while Conscientiousness is lower in boys (Borghuis et al., 2017). A meta-analysis literature portrays that females universally score somewhat higher on all FFM traits and notably higher in Agreeableness (e.g., sympathy and altruism) and Neuroticism (e.g., anxiety and vulnerability) (Schmitt et al., 2016). Similarly, implicit measures of FFM, where participants do not self-report through surveys but are allowed to react to cues, have shown similar results, especially in Agreeableness and Neuroticism (Vianello, Schnabel, Sriram, & Nosek, 2013). Sex and age differences in personality facet traits, seems to replicate cross-culturally, as reported in large samples across 22 nations (Kajonius & Mac Giolla, 2017). Not too many other studies have utilized facets of personality in very large samples.

Note. N = 320,128. The raw scores of the 30 trait facets ranged from 4 (Min) to 20 (Max). N_{Male} = 127,695, N_{Female} = 192,433. 95% CI for effect sizes were ± 0.01 (rounded to two decimals). Bold numbers show above average effects (d ≥ 0.40). A negative Cohen’s d implies that men had lower scores than women.

1 Across 55 cultures, Neuroticism reported higher in females (d = 0.40), Agreeableness (d = 0.15), Conscientiousness (d = 0.12), and Extraversion (d = 0.10), while Openness was inconclusive; using the short version 44-item Big Five Inventory (BF; Schmitt et al., 2008)
based on 92 longitudinal samples (measuring the same person between 10 and 80 years old) showed that Agreeableness and Conscientiousness increase the most in both sexes along the life-span (Roberts, Walton, & Viechtbauer, 2006). Neuroticism on the other hand lowers throughout life, and tends to decrease more in women, while Extraversion and Openness are more ambiguous (Soto, John, Gosling, & Potter, 2011).

3. The present study

In the current study, the main objective was to update and report on scope and size of sex differences in 30 personality facets in a large public sample, using a public domain version of the Five Factor Model (IPIP-NEO-120). Based on existing recommendations, effects around $d \sim 0.40$ (equivalent to $r \sim 0.20$) would be considered notable (Gignac & Szodorai, 2016). The resulting discussion could facilitate our continued understanding as researchers and practitioners (see Hyde, 2014).

4. Method

4.1. Sample and procedure

We used a US sample ($N = 320,128$) surveyed online on five personality trait domains and 30 specific facet traits. The respondents were 19–69 year old, which is the time-frame when traits are known to be stable, and show predictable change throughout age-spans (Briley & Tucker-drob, 2014). The sample consisted of 40% male ($N = 127,695$) and 60% female ($N = 192,433$), with an average age of 28.13 years ($SD = 10.14$). The internet site drew visitors from all walks of life who could have reached the site by word-of-mouth, search engines, or other conceivable channels. Every participant had to actively accept that the survey would be time-consuming (30 min) and that the purpose was scientific research, (also being notified that careless responding would invalidate the worth of the data). The study followed the guidelines from the National Committee for Ethics and did not involve experiments with human subjects. The questionnaire was anonymous, non-sensitive, and volunteered, and is furthermore publically available for researchers (https://osf.io/tbmh5).

One concern was that the unsupervised online collection of data would be characterized by careless answering or deliberate misrepresentations. After removing responses characterized by repetitive patterns ($> 7$ similar answers in a row), the remaining missing data rate was approximately 1%, which was corrected by imputing item means for each variable. The exact details of this procedure are found in Johnson (2005). Due to the very large size of sample, correlations
before and after data cleaning were virtually identical \( (r = 0.99) \).

4.2. FFM facet measurement

IPIP-NEO (Johnson, 2014) was created as an open-source representation of the arguably most used FFM instrument, NEO-PI-R (Costa & McCrae, 1995). The ICC profile correlation between these is almost identical \( (r = 0.98) \) (Maples, Guan, Carter, & Miller, 2014). IPIP-NEO is a product from the collaboration of the International Personality Item Pool (IPIP; Goldberg et al., 2006). In this version, 120 items are measured on a 1 (not very much, almost never) to 5 (very much, almost always) scale, and summarized into facet traits (4 items per facet). Facet traits are then summed and averaged into trait factors (6 facets per trait factor). The five trait factors had high alpha reliabilities ranging from \( \alpha = 0.82 \) (Openness) to 0.90 (Neuroticism and Conscientiousness). On facet level, the mean reliability was \( \alpha = 0.78 \), and only four facets (13%) were slightly below \( \alpha < 0.70 \). See a summary in Table 1.

5. Results

Table 1 summarizes sex differences in the five trait domains as well as 30 specific trait facets, using the effect size of Cohen's \( d \) (mean differences in standard units). Women technically scored higher in all five trait domains; but only notably so in Agreeableness \( (d = 0.58) \) and Neuroticism \( (d = 0.40) \). Trait facets Anxiety (N1), Vulnerability (N6), Emotionality (O3), Altruism (A3), and Sympathy (A6) were large \( (d > 0.50) \). Men showed markedly \( (d > 0.20) \) higher scores only in Excitement-seeking (E5) and Intellect (O5). In summary, almost 50\% (13 out of 30) of the facet traits showed above small effects between sexes, and almost 25\% (7 out of 30) showed above medium effects, while the remaining were trivial in size.

In order to further portray the scope and size of sex differences in facet traits, Fig. 1 illustrates the age-span (cross-sectional) differences in five graphs (Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness). Overall, across all age-groups women scored higher than men in all five trait factors. (The one exception was Extraversion which showed similar levels in both sexes in the 19–29-year-old span). The size of gap between sexes was also fairly uniform in all five trait domains, and Openness stayed along a baseline throughout most of life with some fluctuations. In conclusion, Fig. 1 represents the personality sex differences in one of the most comprehensive FFM facet measurements in the large US public to date.

6. Discussion

The current study showed that almost 50\% of the specific FFM personality trait facets showed above small effects, and almost 25\% above medium effects in sex differences. The most notable difference was seen in the trait domains Neuroticism and Agreeableness. Some specific facets, such as Anxiety (N1) and Sympathy (A6), reported mean effects of over \( d \sim 0.50 \) (Table 1). Interestingly, Neuroticism was, unlike Agreeableness, not uniformly different between sexes across the age-spans, with the largest gap found in the late teens, narrowing and stabilizing first at around 45 years of age. According to a broad evolutionary perspective, this trend seems to coincide with female sexual fertility. In this phase of life, females tend to be more vulnerable than males, in regard to the heightened male sexual aggression (Archer, 2004), while simultaneously investing in pregnancy and caring for infants (Wood & Eagly, 2002). Having women more disposed to anxiety (and empathy), while men are more disposed to assertiveness, may have been an optimal strategy for the propagation of the human species. Certainly, part of the sex gap could also be explained by cultural factors, such as young men not admitting to questionnaire-items assessing neuroticism. However, this explanation may not be supported by other-reports and behavioral observation (Vianello et al., 2013).

Differences in the other trait domains in the FFM were smaller (Openness, Extraversion, and Conscientiousness), and tended to be driven by single specific facets, such as Openness to Emotions \( (d = 0.64) \), Conscientiousness Achievement \( (d = 0.25) \), and Extraversion Activity \( (d = 0.24) \). Overall, these sex differences in the present US sample (Table 1) aligned well with the now almost 20-year-old landmark findings in the original FFM NEO-PI-R model (Costa Jr et al., 2001). Comparing the sex gap in facet traits in US adults in our present study with Costa Jr et al. (2001) showed no reversed effects, while a few (e.g., Friendliness, Gregariousness, Trust, and Self-efficacy) had dropped to trivial levels. However, even more traits showed increased sex gaps, which may be implicated by the thesis that the more progressive a society becomes, the greater the sex differences in personality (Schmitt et al., 2008; Stoeet & Geary, 2018).

6.1. Interpreting differences

It is clear that there is a scope of sex differences in personality. However, perhaps even more interesting is how can we understand the size of differences? The overall effect size in the five trait domains was generally speaking 1/4 of a standard deviation. One way to understand an effect of \( d = 0.25 \) standard scores would be to pick a random man and women from a population and have a 56\% probability of the one scoring higher than the other (Cohen, 1992). This is about the similar odds a casino needs to ensure a practical profit advantage. In the present study, several facets differed by more than double that, 1/2 standard deviation \( (d = 0.50) \). Such effect sizes would have societal implications, for instance in how men and women perceive one another, or what sex is mostly represented in certain educational programs or occupations. 1/2 standard deviation is equivalent to approximately 70\% of the one sex scoring higher than the mean in the other (See http://rpsychologist.com/d3/cohen-d/), which could help explain overrepresentations in certain life careers due to varying personal interests (e.g., engineering vs nursing; See Stoeet & Geary, 2018). The implication at large for society is that certain behaviors and choices will be considered as feminine or masculine by most. Using a practical analogy in the form of a flipped coin – An effect of \( d = 0.50 \) would imply that heads come up 63/37 instead of the normal and expected 50/50 (Ozzer, 1985). Again, such odds seem likely to translate into societal impact.

In addition to these statistical enlightenments, it is noteworthy to consider that none of the effect sizes have been disattenuated. The always present random error in personality scales likely implies that we tend to underestimate the real sex differences. Furthermore, one school of thought contends that sex differences generally are underestimated due to the multidimensional nature of personality. The argument is that since many traits are most often experienced simultaneously they should not be analyzed separately, trait by trait, as scientists tend to do (Del Giudice et al., 2012). As an illustration, when we estimate the size of any object we may encounter in everyday life, we naturally consider many dimensions simultaneously, such as height, circumference, and constitution, not only one separately at a time. Similarly, we do not experience typical male or female personality traits one by one, but instead as one whole experience. When aggregating and analyzing differences in mating-strategies between sexes in this way, this has resulted in a very large global effect sizes of Mahalanobis \( D = 2.71 \) (Del Giudice et al., 2012). Such an effect of almost 3 standard deviations

---

\(^2\) See visual graphs on Cohen’s \( d \) calculations and percentage of differences between groups at http://rpsychologist.com/d3/cohen-d/.

---

P.J. Kajonius, J. Johnson

Personality and Individual Differences 129 (2018) 126–130
would transfer into 99.7% of men scoring higher than the average woman, at least when it comes to sexual behaviors. This new way of regarding sex differences is under criticism and ongoing debate (Hyde, 2014).

7. Concluding thoughts

Particularly two trait domains seem to be of driving interest when it comes to sex differences; Narcissism and Agreeableness. From a more general scientific viewpoint, it is critical to constantly reaffirm and extend foundations of psychology, especially in the wide-reaching wake of the recognized replication crisis within psychology (Anderson et al., 2016). This study has attempted to heed this by confirming and expanding on the scope and size of sex differences in personality with one of the largest public samples to date.

References


