DOCUMENT RESUME

ED 383 759 TM 023 289

AUTHOR Guo, Shuqin; And Others
TITLE Sex Differences in Personality: A Meta-Analysis Based on "Big Five" Factors.
PUB DATE Apr 95
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adults; *Effect Size; *Females; Males; Meta Analysis; *Personality; Personality Measures; Psychometrics; *Sex Differences
IDENTIFIERS *Big Five Markers; *Self Report Measures

ABSTRACT

Analyses of sex differences in personality have been reported only in particular scales or individual domains. In this study, a psychometric meta-analysis of adult sex differences in self-reported personality was conducted based on each of the "big five" factors. Examination of 76 studies (1,057 separate effect sizes) on 35 personality scales showed that women's scores were significantly higher than men's on measures of Neuroticism and Agreeableness. This finding raises the issue of whether the observed differences are due to differences in self-report or genuine differences in personality. Three tables are included. (Contains eight references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Guo Shuqin
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED 383 759

Sex Differences in Personality: A Meta-Analysis Based on "Big Five" Factors

By

Shuqin Guo
Xiaoping Wang
Thomas Rocklin

The University of Iowa
Division of Psychological and
Quantitative Foundations
College of Education
Iowa City, Iowa 52242

American Educational Research Association
San Francisco
April 1995

CE 069 221

BEST COPY AVAILABLE

Abstract

Analyses of sex differences in personality have been reported only in particular scales or individual domains. In this study, a psychometric meta-analysis of adult sex differences in self-reported personality was conducted based on each of the "big five" factors. Women's scores were significantly higher than men's on measures of Neuroticism and agreeableness. This finding raises the issue of whether the observed differences are due to differences in self report or genuine differences in personality.

Sex differences in Personality: A Meter-Analysis Based on "Big Five" Factors

Introduction

Although sex differences in mean scores on particular personality scales or domains have sometimes been reported (e. g., Hembree, 1988) no comprehensive study of the scores of men and women on a wide range of operationalizations of personality constructs has been reported. The most recent research on a broader range of personality scales was done by Alan Feingold (1994), who reported gender differences in nine facets. In this study, we report a psychometric meta-analysis (Hunter and Schmidt, 1990) of sex differences on the "big five' factors in personality as measured by self report scales.

There is a growing consensus among personality researchers that the broad domain of personality (as measured by self report scales and peer ratings) can be characterized by five dimensions: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness (John, 1990). These five factors have been demonstrated cross a wide range of subjects and several investigators (Digman & Takemoto-Chock, 1981). Although many personality inventories (e.g., the Eysenck Personality Questionnaire, Eysench & Eysenck, 1975) assess only a subset of the "big five," in general each of the traits measured by the major personality scales can be subsumed under one of the big five factors.

Analyses of sex differences in scores on measures of cognitive abilities (e.g., Hyde & Linn, 1986) have been much more common than analyses of sex differences in scores on measures of personality. This may be due, in part, to the relatively less ambiguous interpretation of differences in maximal performance tests as opposed to typical performance tests such as personality scales. In particular, differences in scores on personality tests may reflect either genuine differences in personality or differences in the willingness or ability to respond to questions about personality. The extent to which this distinction is a concern depends on the extent to which personality scales show differential validity for men and women. We will return to this issue in the discussion of our results.

Objective

The purpose of the research reported here is to estimate the magnitude of sex differences in personality scores. In particular, the evidence for sex differences in measures of each of the big five factors of personality will be separately examined.

Methods

Literature Search

A preliminary data base was generated by conducting a computer search of Eric, PsychLit, and the Social Science Citation Index from 1967 through 1992. A personality inventory manual search was conducted as well.

This search procedure located 76 published studies or standardized norms (yielding 1057 separate effect sizes) on 35 different personality scales. Studies were included in this research if they met the following criteria:

(1) Subjects were drawn from normal adult American population no younger than sixteen.

(2) Studies had both male and female subjects with similar sample sizes.

(3) Group means, standard deviations, number of subjects were reported. Where available, reliability coefficients (test-retest or internal reliability) were also collected.

Coding the Variables

Barrick and Mount (1991) had personality scales from various personality inventories rated by six psychologists into six dimensions. The raters were instructed to select a dimension as being representative of a certain personality scale. If over half of the raters agreed with the selection for that scale, it was categorized under that dimension. The six dimensional categories are as follows:

(1) Neuroticism/Emotional stability (High score means Neuroticism).

(2) Extraversion/Introversion (High score means Extraversion).

(3) Openness to Experience (High score means imagination, curiosity, originality, broadmindedness, and so on).

(4) Agreeableness (high score means sympathetic, softhearted and so on).

(5) Conscientiousness (high score means hard working, organized, and so on).

(6) Other (scales that could not be classified as representing one of the big five factors).

Where necessary, scale scores were reversed to make the direction of scoring consistent within a category.

Analyses

Effect sizes (ES or d) in this meta-analysis indicate gender differences in central tendency, the differences between means of males and females divided by the pooled standard deviation. In this study, formulas from Hunter and Schmidt's book Methods of Meta-Analysis (1990) were used to calculate the effect size and the variance.

$$ES = \frac{(Mm - Mf)}{\left[(SDm^2 * Nm + SDf^2 * Nf)/(Nm + Nf)\right]^{1/2}}$$

(M_m and m_f represent the means of males and females, N_m and N_f are the sample sizes of males and females, SD_m and SD_f are the standard deviation of males and females).

In current study, both uncorreccted and corrected analyses were separately performed in measures of each of the big five factors for personality. The data were analyzed separately for scales for which internal consistency reliability or test-retest estimates were available.

For the uncorrected analysis, mean effect sizes can be obtained with a sum of $ES*(N_m+N_F)$ divided by overall number of subjects.

$$\overline{ES}=\sum[ES_i*(N_m+N_f)]/N_{total}$$

where N_{total} equals to the total number of all the cases
Standard deviation of effect sizes can be obtained from the variance of population effect size Var(δ).

$$Var(\delta) = Var(d) - Var(e)$$

The formulas for Var(d) and Var(e) are

BEST COPY AVAILABLE

$$Var(d)=\sum[(ES_i-\overline{ES})^2*(N_m+N_f)]/N_{total}$$

$$Var(e)=(4/N)*(1+d^2/8)$$

where N equals to the average number of sample size for each dimension.

For large sample size population mean (δ) is similar to sample mean effect size as the sampling error is close to zero.

For corrected analysis, the effect sizes are corrected with the reliability of either test-retest or internal consistency, whichever is available. True population effect size can be obtained with the formula

$$\delta = ES/a$$

where a equals to the square root of the reliability.

Results

Results from the Uncorrected Analyses

Table 1 shows the results from the uncorrected analyses. Negative effect size means were found in Neuroticism (ES =-.25), Agreeableness (ES = -.34), and Conscientiousness (ES=-.10), which indicate higher scores for women than for men. Positive effect size means with a male advantage were found in Openness (ES=.13) and Extraversion (ES=.10). The standard errors of the effect sizes ranged from .04 to .08.

Results from the Corrected Analyses

Table 2 and 3 show the results for the corrected analyses. When test-retest reliability was used for correction, the mean is -.26 for Neuroticism, -.41 for Agreeableness, -.09 for Conscientiousness , .19 for Extraversion, and .18 for openness. The range of standard errors are from .06 to .11.

When internal consistency reliability was used for correction, only the values of the means of effect size in Neuroticism (ES=-.32) and Agreeableness (ES=.-36) were increased from the uncorrected analyses. The values of the other three means were all dropped in some degree. This might be explained by the decrease of the number of cases that were available with internal consistency reliability and the low reliability. The standard errors for the analyses corrected with internal consistence reliability ranged in between .06 and .08.

For these two kinds of correction, Neuroticism and Agreeableness have higher reliability than Openness, Extraversion and Conscientiousness do. Neuroticism and Agreeableness tended to have smaller standard errors of effect sizes than the other three in both uncorrected and corrected situation. Data with small standard error can develop a smaller confidence interval to give an accurate estimation. The categories with small standard error and high reliability also represent a better defined domain. Thus Openness and Conscientiousness may represent a broad characteristic.

Discussion

Results of meta-analysis are informative to understanding differences between natural groups. In this study it is found that women scored considerably higher on scores of Neuroticism and Agreeableness and somewhat higher on Conscientiousness, and men, on the other hand, scored higher on Openness and Extraversion. Generally speaking, intervals of confidence are computed to see whether an ES is significant or not. As the sample size of the studies in this research are very large (the average sample size was greater than 1000), none of the formulas is flawless to calculate the standard deviation of the population effect size. Cohen (1977) provides rough guidelines of ES=.20 (small effect), ES=.50 (medium effect), ES=.80 (large effect), with the caveat that it is better to obtain these standards for comparison from the professional literature than to use these somewhat arbitrary guidelines. Hunter and Schmidt mentioned in their book that most textbooks of social science claims very rare that an effect size exceeds .40. Rossi and Wright (1977) suggested that .50 standard deviation improvement is considered to a conventional measure of practical significance. Borrowing these ideas it is safe to say that there are significant sex differences on Neuroticism and Agreeableness with scores in favor of women. For Extraversion men scored higher than women, but the significance is at the boarder line. There are no significant sex differences for Openness and Conscientiousness.

Errors of measurement is a serious issue in meta-analysis. In this study standard effect sizes are used, with each effect size divided by the pooled within group standard deviation. Large errors of measurement can increase population variance and thus reduce effect size. For our study, the sample size for each study is very large and there are plenty studies for each dimension. Measurement errors does not alter the tendency of the sex differences in personality.

Being able to delineate sex differences in personality within the big five is a first step in understanding the sources of those differences. This finding raises the issue of whether the observed differences are due to differences in self report or genuine differences in personality. Further research exploring the meaning of these differences is justified.

References

Barrick, M. R. & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. Personnel Psychology, 44, 1-25.

Eysenck, H. J. & Eysenck, S. B. G. (1975). Manual for the Eysenck Personality Questionnaire. San Diego: Education and Industrial Testing Service.

Feingold, A. (1994). Gender Differences in Personality: A Meta-Analysis. Psychological Bulletin, 116 (3) 429-456.

Hembree, R. (1988). Correlates, causes, effect, and treatment of test anxiety. Review of Educational Research, 58, 44-47.

Hunter, J. E. & Schmidt, F. L. (1990). Methods of meta-analysis: Correlating error and bias in research findings. New bury Park, CA: Sage.

Hyde, J. S. & Linn, M. C. (1986). The Psychology of Gender: Advances Through Meta-Analysis. The Johns Hopkins University Press, Battimore.

John, O. P. (1990). The "big five' factor taxonomy: Dimensions of personality in natural language and in questionnaires. In L. A. Pervin(eds.), Handbook of Personality Theory and Research New York: Guilford.

Wolf, F. M. (1986). Meta-Analysis: Quantitative methods for research synthesis. New bury Park, CA: Sage.

Table 1: Means and Variances of Effect Sizes for Uncorrected Analysis

Dimension	Total Number	Total Cases	Mean Effect Size	Var(d)	Var(e)	Var(δ)
Neuroticism	290,441	231	-.25	.14	,003	.15
Extraversion	227,559	156	.10	.10	.003	.087
Openness to Experience	275,184	138	.12	.20	.002	.198
Agreeableness	309,266	226	-.34	.10	.003	.097
Conscientiousness	209,980	127	-.10	.08	.002	.078

Table 2: Means and Variances of Effect Sizes for the Analysis Corrected with Test-Retest Reliability

Dimension	Total Number	Total Cases	Mean Effect Size	Var(d)	Var(e)	Var(δ)
Neuroticism	196,451	131	-.26	.21	.003	.227
Extraversion	155,419	83	.19	.13	.002	.127
Openness to Experience	202,797	80	.18	.22	.002	.218
Agreeableness	215,892	108	-.41	.11	.002	.108
Conscientiousness	120,916	53	-.09	.14	.002	.138

BEST COPY AVAILABLE

Table 3: Means and Variances of Effect Sizes for analysis corrected with Internal Consistency reliability

Dimension	Total Number	Total Cases	Mean Effect Size	Var(d)	Var(e)	Var(δ)
Neuroticism	145,465	108	-.32	.10	.003	.097
Extraversion	128,770	85	-.03	.14	.002	.138
Openness to Experience	111,255	85	.01	.20	.004	.196
Agreeableness	179,181	161	-.36	.13	.003	.127
Conscientiousness	141,294	103	-.03	.10	.002	.098