MANOVA

**Description:**

MANOVA stands for a multivariate analysis of variance. Specifically, MANOVA is most commonly used when you want to examine the relationship of one to many independent variables to two or more dependent variables. Regular ANOVA is for one to many independent variables with *one* dependent variable. MANOVA lets you combine dependent variables to look at overall group differences on a combined set of variables. Then you can break down the variables and see which individual dependent variable contributes to group differences the most.

**Definitions/Abbreviations:**

IV – independent variable. This variable *has* to be a dichotomous variable. You can put people into groups based on any category (gender, handedness) or your experimental manipulation (instructions versus no instructions).

Between subjects – between subjects variables are used here for independent variables. Between subjects variables are variables with different groups or labels and cannot have continuous values. MANOVA is a between subjects ANOVA with several dependent variables.

DV – dependent variable. The dependent variable *needs* to be a continuous variable or another type of analysis might work better (see log regression). Your dependent variable should be the measurement you took in your study or what information you are expecting to see changed over groups.

DV combinations – Wilk’s Lambda, Roy’s Largest Root, Hotellings Trace, Pillia’s – these are all listed in the multivariate test. They are different ways to combine the DVs in such a way that creates large group differences on your IVs. Think of these as “giant means” for your DV, if you could create one mean of all the DVs such that your groups were maximally different. The most commonly used is Wilk’s Lamba.

**The process:**

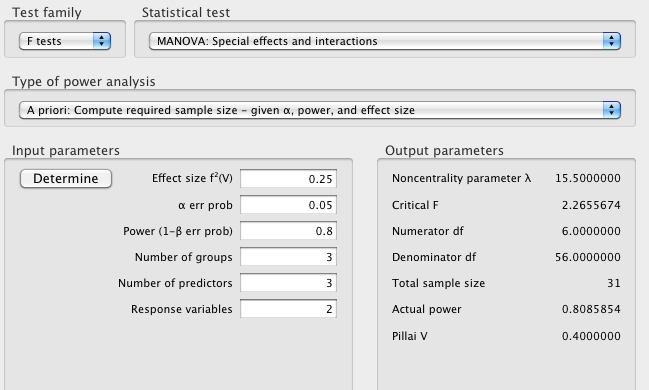
**Power:**

You want to check how many people you need to run (or alternatively, how many more people you need in a study).

Using G\*Power, finding ideal participant numbers is fairly easy. Set up options:

* Test family: F-tests
* Statistical Test: MANOVA: Special Effects and Interaction (for all between subjects).
* Type of power analysis: A priori (most common)
* Effect size f: either guess at an effect size based on research, use a small effect size for good measure, or after a couple subjects run a prelim MANOVA and use the current effect size. (You can click determine to convert eta squared to f).
* Alpha = .05
* Power = .80
* Number of groups = Number of conditions
* Number of predictors = number of IVs
* Number of response variables = number of DVs

Hit calculate for the number of participants needed.



Note: Pillai V will be part of the output (under value for Pillai) but it is still easiest to calculate from eta squared or estimate based on the numbers it shows you for small/medium/large when you hover over the effect size box.

**Assumptions:**

Outliers:

* Univariate – outliers only on one of the DVs. You want to check z-scores for people who are more than 3 (3 or -3) away from the mean. This analysis will check each DV separately.
* Multivariate – outliers on the combination of the DVs. This procedure uses Mahalanobis distance to make sure that people do not have a strange combination of answers all of the dependent variables.
  + You can do both of these or just Mahalanobis. You may not want to eliminate people who are a univariate outlier on one variable, but you really should eliminate people who are multivariate outliers (especially since this test is multivariate!)

Multicollinearity: You want to check your DVs by using a correlation to make sure they do not overlap too much. If they overlap a great deal, then you want to use only one of them or combine them before running. Look for variables with r > .9.

Linearity: Linearity between the DVs is a very important issue because the combinations made of the giant DV are linear. You can check for this value using a fake regression or bivariate scatterplot.

Normality:

* Univariate – you want DVs to be normally distributed by themselves. You can check this information through frequencies and asking for a histogram. Non-normal distributions also have skew and kurtosis values over 3/-3.
* Multivariate – you also want the DV combinations to be normally distributed. You can check for multivariate normality by running a fake regression and asking for a histogram of the residuals.

Homogeneity: the variance of the groups from your IV need to be equal across all of the DVs. You can check this information with a residual plot from your fake regression (you do not want raining or an unequal spread of the dots around 0). You can also use Box’s M test of homogeneity – you *do not* want p<.001.

# Complete Example

Researchers have measured participants on their femininity and masculinity and want to know how those two variables affect a range of dependent measures. They measured self-esteem, attitude about women’s roles, and neuroticism to see if there were differences across femininity and masculinity scores.

**IVS:**

Femininity scale (low versus high)

Masculinity scale (low versus high)

**DVS:**

Self esteem

Attitude toward the role of women

Neuroticism

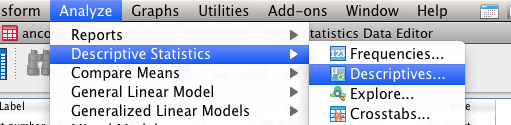
**Research question:** We want to see if the different feminine-masculine groups will score differently on a combination of DVs. So we might see if they differ on “personal factors”, which is the combination of several of our DVs.

**Assumption Checks:**

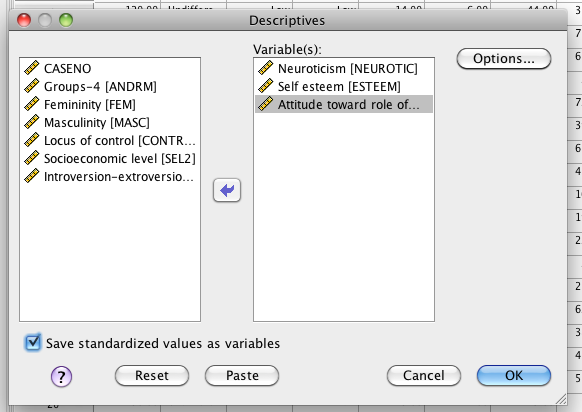
Accuracy and missing data are same as data screening notes. Be sure to check if the data is accurate and has no missing pieces first.

Univariate outliers:

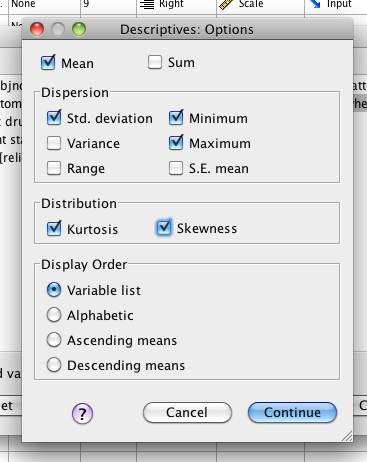
1. Analyze > Descriptive Statistics > Descriptives



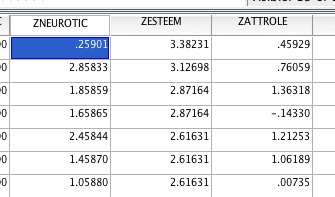
1. Move over all your DVs into the variables box on the right.
2. Hit save standardized values as variables.



1. Hit options > skew and kurtosis.



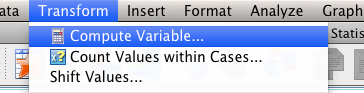
1. Look for univariate outliers.
2. Go back to the data set. You should see new variables created for your three different DVs.



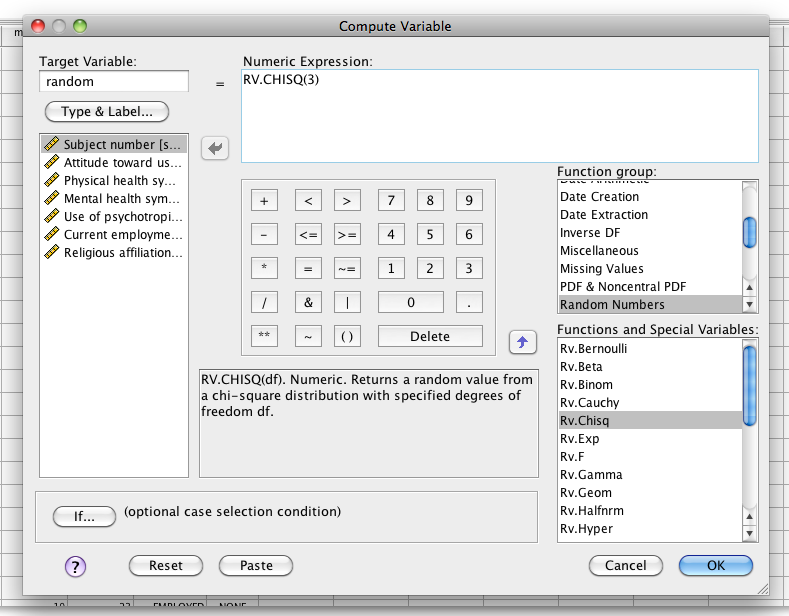
1. Sort those variables one at a time – look for values over 3 and -3 (right click sort). Here we see one for self esteem.
2. Check for multivariate outliers before you eliminate univariate outliers. You will want to eliminate multivariate outliers more than you would univariate outliers.

Multivariate outliers:

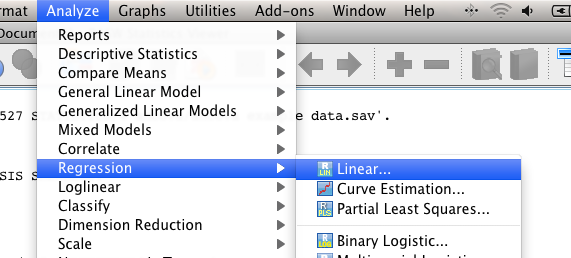
1. Create a fake variable to use for your fake regression.
2. Transform > compute variable.



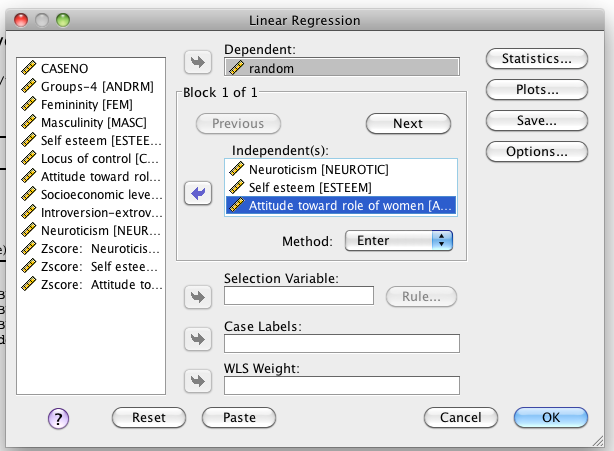
1. Call the target variable “random”. Use one of the random functions to create your random variable (I’ve used ChiSquare). The functions often have a ? for you to fill in a number (note this is a repeat picture).



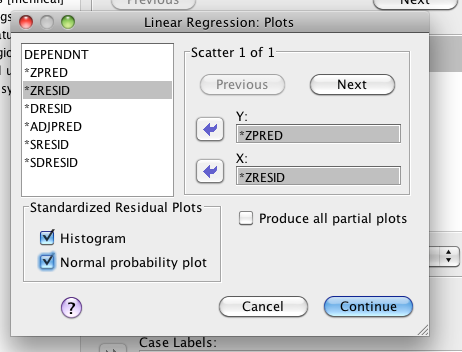
1. Hit “ok” and it will create a new random variable for you.
2. Run a fake regression. Analyze > Regression > Linear.



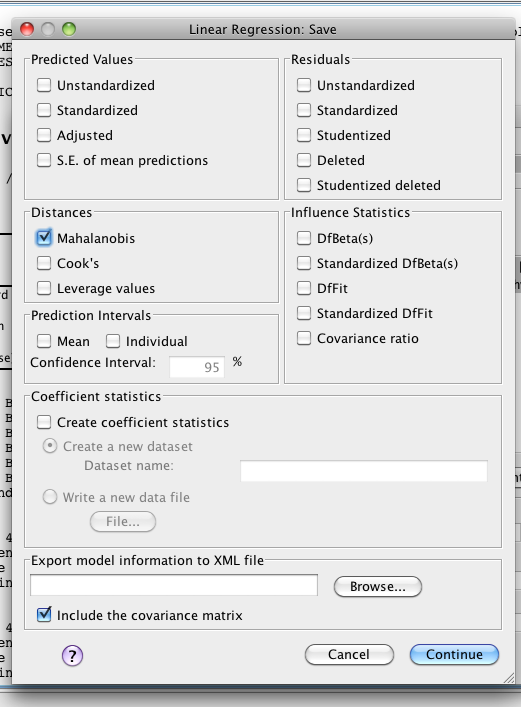
1. Move your random variable into the DV box. Put your DVs into the independents box.



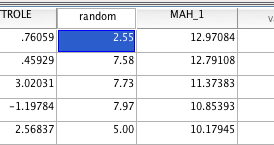
1. Hit Plots.



1. Put zpred in Y and zresid in X. This section will create the residual plots for other assumptions checks. Check histogram and normal probability plot.
2. Hit Save.



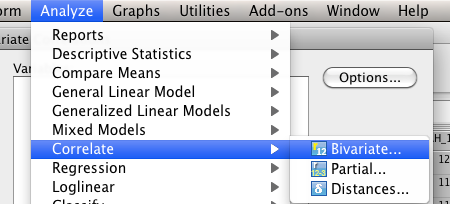
1. Check Mahalanobis – it will create another new variable for you to check.
2. Go back to the data set, and sort cases descending for Mahalanobis.
   1. You will need a cut off value to understand these scores.
   2. The cut off value is chi-square with degrees of freedom = number of variables, p<.001.
   3. Here we have 3 variables (3 different DVs), p<.001, chi square = 16.27 cut off score.
   4. You will look for people who have a Mahalanobis score greater than > 16.27.



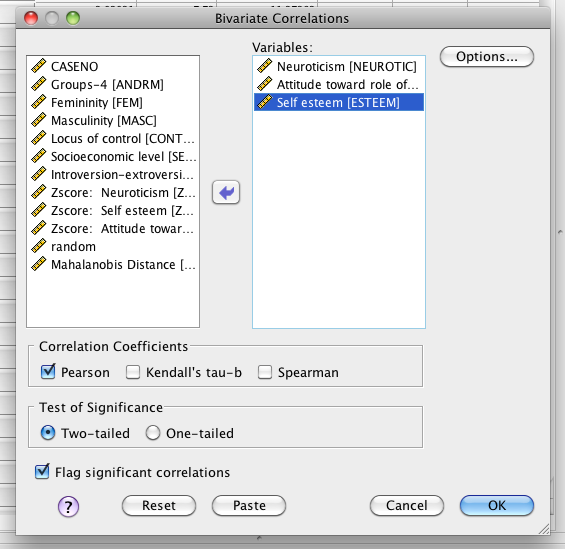
1. We do not have multivariate outliers. I am not going to exclude the univariate outliers because they were not a multivariate outlier.

Multicollinearity – you want to make sure that your DVs are not too highly correlated or you will run into trouble when it tries to combine them. Note that you will also lose power if your DVs are too uncorrelated (i.e. *r* < .10).

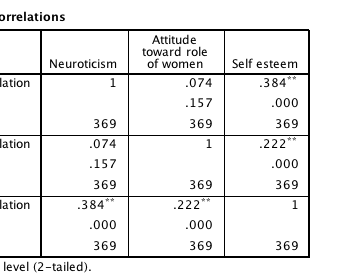
1. Analyze > correlate > bivariate.



1. Move all your DVs to the right hand side and hit ok.



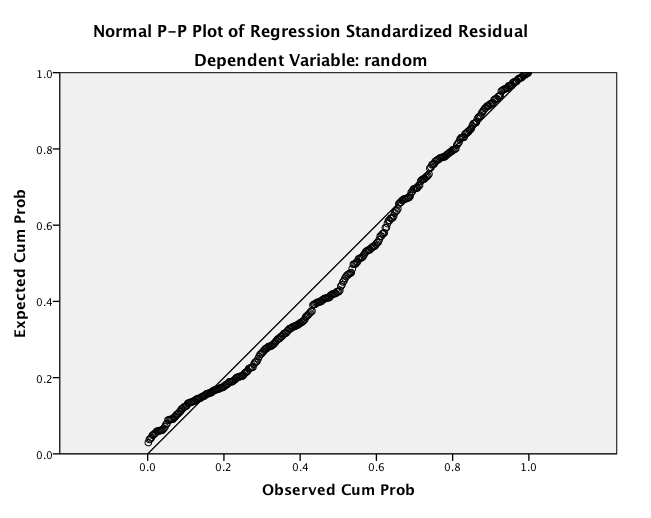
1. You want to make sure no correlations are over .9.



1. None of these are greater than .9, so we would be ok.

Linearity:

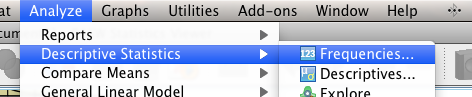
1. You can check linearity with the fake regression you’ve already run.
   1. IF YOU DELETED OUTLIERS: run the regression again, just as it describes above.
   2. IF YOU DID NOT DELETE OUTLIERS: you can use the output you already have.
2. Scroll down to the Normality Probability Plot.



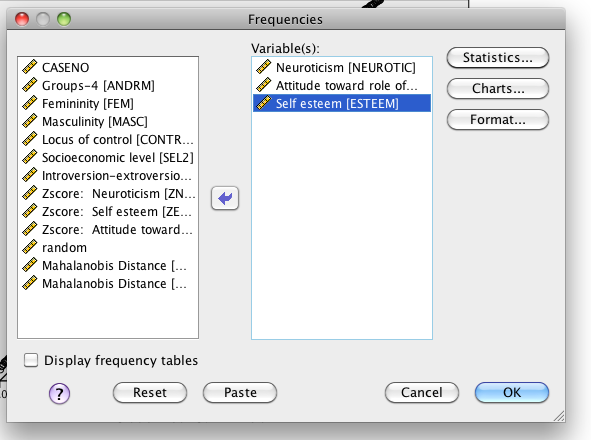
1. You want the dots to follow along the line moderately closely. This picture is very linear.

Normality:

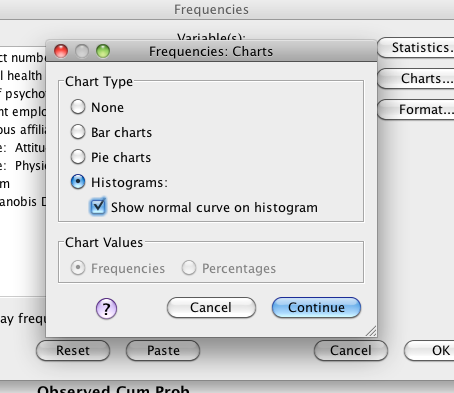
1. Univariate normality:
   1. You will need to get individual histograms of your DVs.
   2. Analyze > Descriptive Statistics > Frequencies.



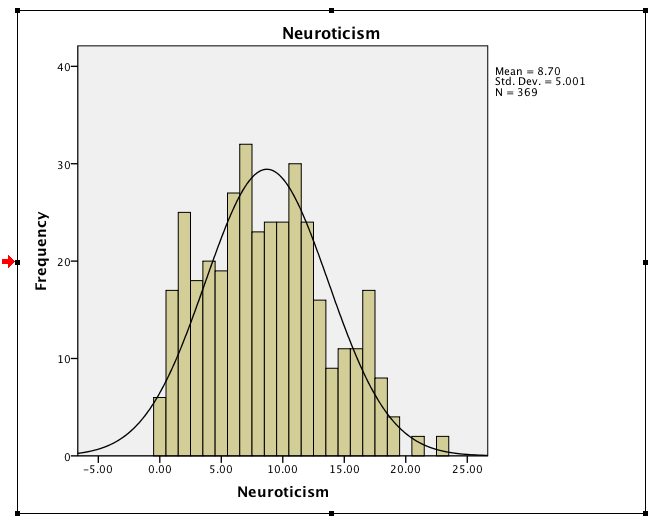
* 1. Move over the CV and DV into the right hand box. I turned off frequency tables because I didn’t want them.



* 1. Hit charts. Pick Histogram and Normal Curve.

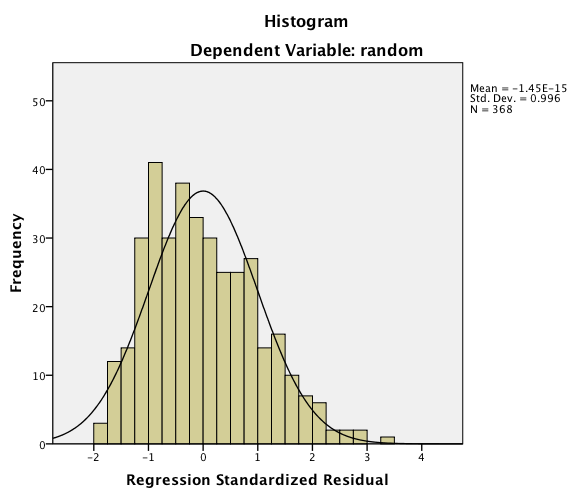


* 1. Hit continue and Ok.
  2. Now you’ll want to check the charts to make sure they look ok.



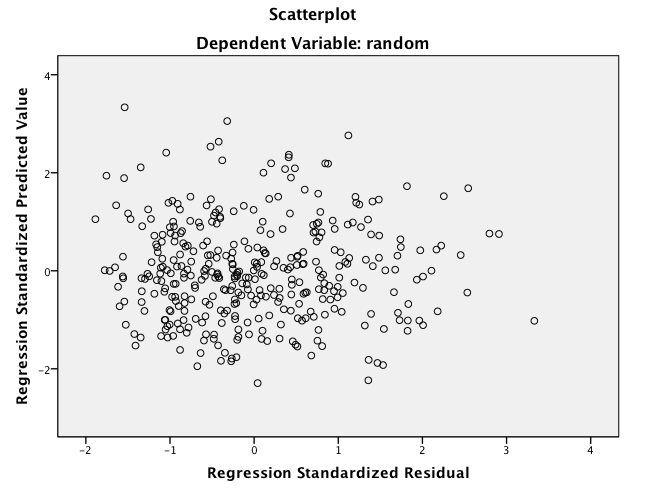
* 1. These look fine. If you aren’t sure you can run frequencies (see above) and see if the values are 3 or -3 for skew and kurtosis.

1. Multivariate Normality: use the histogram created from your fake regression to check DV combination.



* 1. This histogram looks fairly normal.

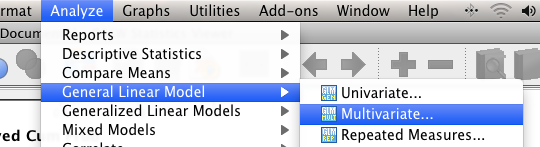
Homogeneity: you can check these assumptions using the residual plot from the fake regression you ran earlier.



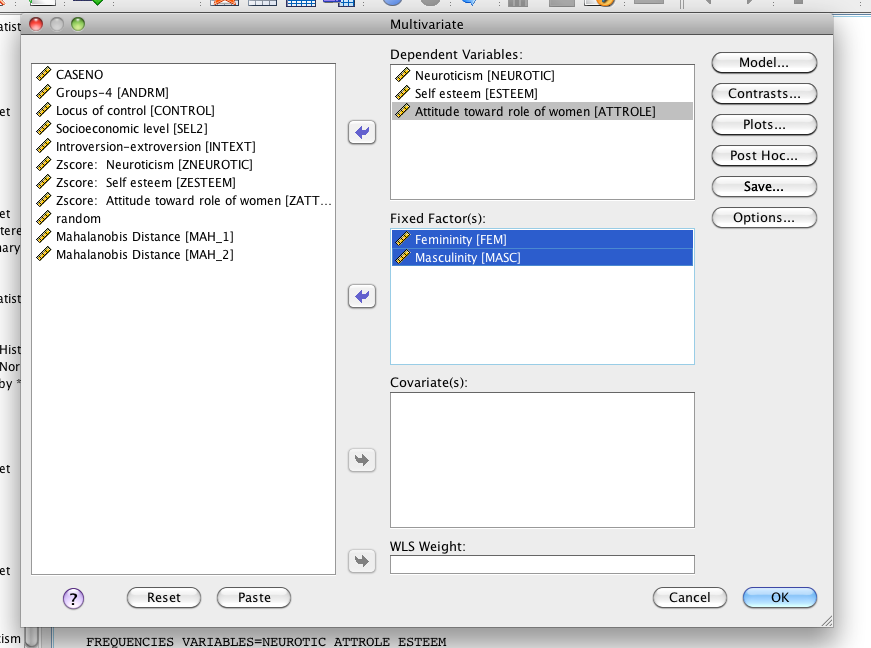
1. What to look for:
   1. Draw a line at 0.
   2. Homogeneity – is the spread above that line the same as below that line?
      1. These lines are complete lines.
      2. So it goes from 0 to 4 above and 0 to -3 below.
      3. You *do not* want a very large spread on one side and a small spread on the other side.
      4. If you encounter this problem – check Box’s and Levene’s when you run the data.

**Running the Analysis:**

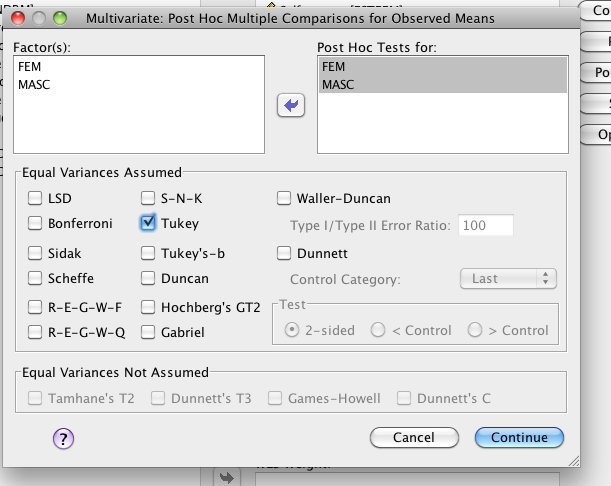
1. Analyze > General Linear Model > Multivariate



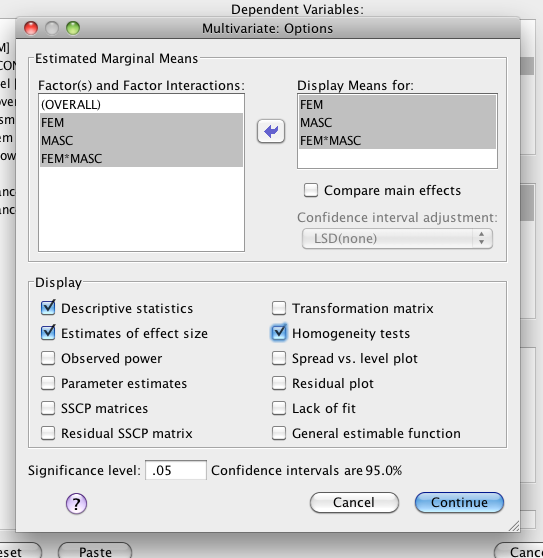
1. Put your DVs (attrole, self-esteem, neuroticism) into the Dependents Box. Put your IVs (fem and masc) in the fixed factors box.



1. Hit Post Hoc. Move over your IVs you want a post hoc test for and ask for Tukey. If you only have two levels, you can still ask for this information, but it will not run. It helps to get into the habit, so you don’t forget when you have 3 or more levels.



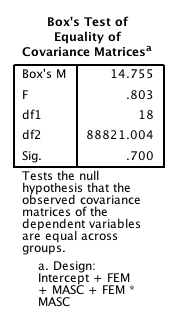
1. Hit options. Move over your IVs to the right hand side. Ask for effect size, descriptive statistics and homogeneity (Box’s M and Levene’s).



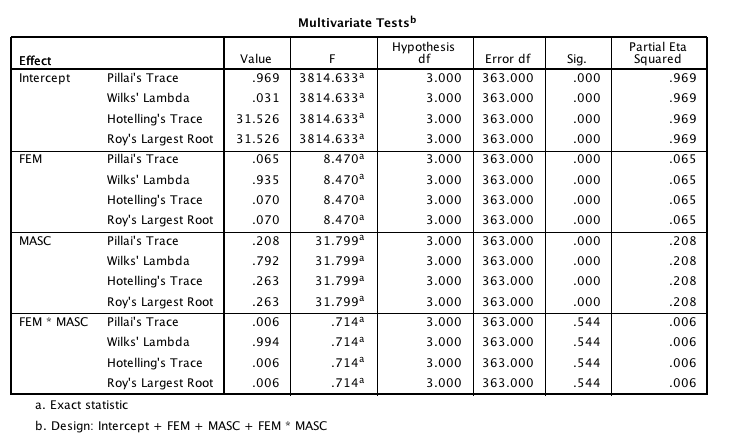
1. Hit continue and ok.

Reading the output:

1. Homogeneity in Box’s M:
   1. If you were not sure about your homogeneity test with the residual plot, then you can check Box’s M for multivariate homogeneity. You *do not* want *p* < .001. Here we are ok because Box’s M *p* = .70.
   2. Why not just get this number every time? Why are we doing it the fake regression way?
      1. You can only get Box’s M appropriately (i.e. the math is right) for multiple DV situations with categorical IVs. We learned the data screening in regression so we can use those same rules for analyses with continuous IVs (i.e. regression, EFA).

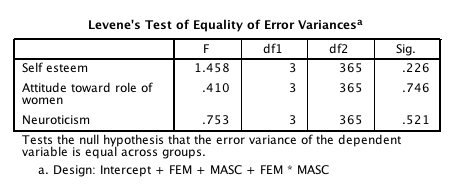


1. Multivariate test box – this box will give you the information for the MANOVA for each main effect (fem, masc) and the interaction (fem\*masc).

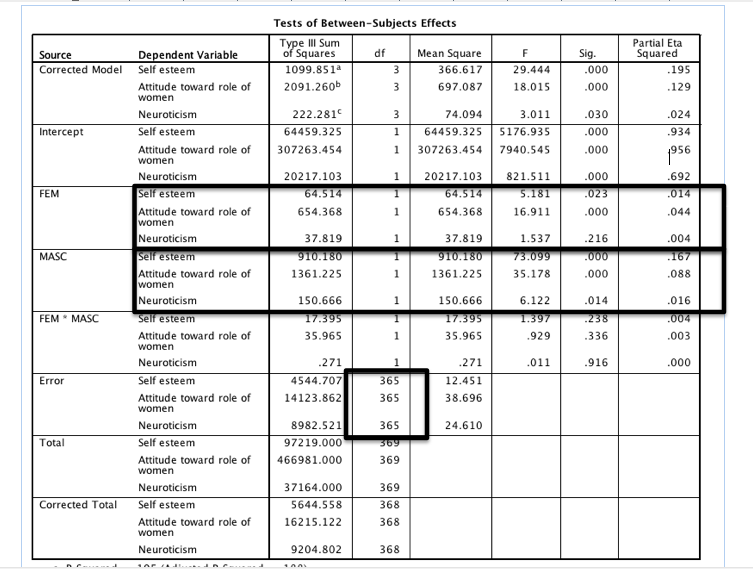


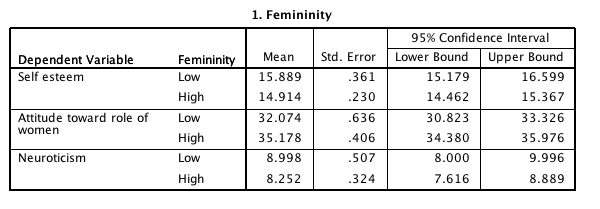
* 1. Main effect Femininity: (don’t forget the DF go across in multivariate tests and not down).
     1. There was a significant effect of Femininity, *F*(3, 363) = 8.47, *p* < .001,*ηp2* = .07.
     2. Because there was a significant effect, you will also want to look at the regular ANOVA for FEM.
  2. Main effect Masculinity: There was a significant effect of masculinity, *F*(3,363) = 31.80, *p* < .001, *ηp2*= .21.
     1. Because there was a significant effect, you will also want to look at the regular ANOVA for Masculinity.
  3. Interaction: There was not a significant interaction between masculinity and femininity, *F*(3,363) = .71, *p* = .54, *ηp2* = .01.

1. Levene’s Test – Levene’s is for univariate homogeneity and would be useful if your Box’s M was significant (p < .001). You could then tell what DV is the problem. Since we met multivariate homogeneity, we can move on.

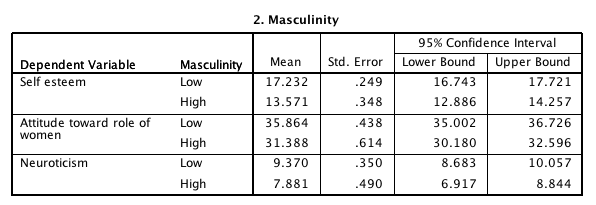


1. Regular ANOVA / Between Subjects Box – after you run a MANOVA, the individual ANOVAs run for each IV-DV combination. IF your main effects/interaction were significant, you will look here as your second step. This box will tell you which DV(s) your group differences were found on. If your effects were not significant, you would ignore their counterpart in this box.



1. Main effect of Femininity:
   * 1. Self esteem significant main effect, *F*(1, 365) = 5.18, *p* = .02, *ηp2*= .01.
     2. Attitude of women’s role significant main effect: *F*(1, 365) = 16.91, *p* < .001, *ηp2*= .04.
     3. Neuroticism did not have significant main effect: *F*(1, 365) = 1.54, *p* = .22, *ηp2* = .004.
2. Main effects of Masculinity:
   * 1. Self esteem significant main effect, *F*(1, 365) = 73.10, *p* < .001, *ηp2*= .17.
     2. Attitude of women’s role significant main effect: *F*(1, 365) = 35.18, *p* < .001, *ηp2*= .09.
     3. Neuroticism significant main effect: *F*(1, 365) = 6.12, *p* = .01, *ηp2*= .02.
3. Post Hoc Tests – you only want to do post hoc tests for the significant effects you found.
   1. For main effects – you can run a Tukey Post Hoc test (report as you would with regular ANOVA).
   2. For main effects with only 2 levels – just compare the means.
   3. For interactions – you can run independent t-tests on the conditions (remember that conditions are the individual cell combinations aka Fem HIGH Masc HIGH).
4. Post Hoc Means:
   1. Femininity means: you’ll want to talk about the means and SE for each significant DV.

b. Masculinity means: you’ll want to talk about the means and SE for each significant DV.



**Effect size:**

**Use** *ηp2* for the MANOVA and ANOVA results, and use MOTE to calculate the *d* values for the post hoc tests when necessary (use independent t page) by entering the means and standard deviations. You also want to make sure to enter the right *n* values for each group.

**Charts:**

Graphs > chart builder

Pick bar graph on the left side.

Since we have two variables, we are going to use a clustered bar graph.

Put one IV in the x-axis, one IV in the set color option. Usually you want to pair together the options that you used in your post hoc (i.e. here I would put my categories together because that’s what I compared.). Put the DV in the y-axis.

Be sure to use error bars, x and y axis labels. See figure below.

**Note: You will have to make a separate chart for each DV.**

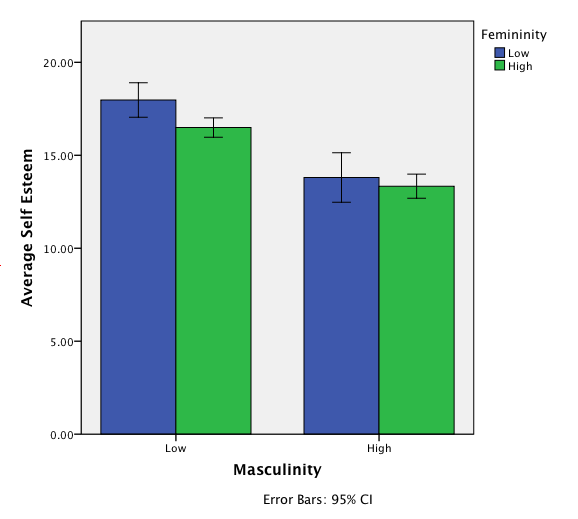
**Example Write Up:**

**Results**

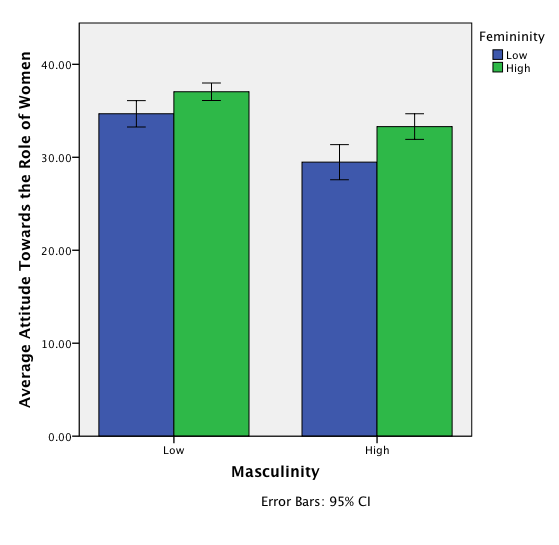
Prior to analysis, data were screened for missing data and outliers. Several participants were identified as univariate outliers, but no cases were multivariate outliers (using Mahalanobis distance). All participants were retained for this analysis. Data were found to be multivariate normal, linear, and homogeneity was met (Box’s M *p* = .70). A 2 X 3 between subjects MANOVA was analyzed with femininity (low, high) and masculinity (low, high) predicting neuroticism, self-esteem, and attitude toward the role of women. Figures 1, 2, and 3 depict the average scores for the conditions on the dependent variables.

Significant multivariate main effects were found for femininity (*F*(3, 363) = 8.47, *p* < .001, *ηp2* = .07) and masculinity (*F*(3,363) = 31.80, *p* < .001, *ηp2*= .21), but not for the interaction between femininity and masculinity (*F*(3,363) = .71, *p* = .54, *ηp2* = .01). Univariate ANOVAs were used to examine individual dependent variable contributions to main effects. Femininity scores showed a significant difference in self-esteem scores (*F*(1, 365) = 5.18, *p* = .02, *ηp2*= .01), where low femininity participants (*M* = 15.89, *SE* = .36) scored higher than high femininity participants (*M* = 14.91, *SE* = .23). Attitudes toward the role of women showed the opposite effect (*F*(1, 365) = 16.91, *p* <.001, *ηp2*= .04), where low femininity (*M* = 32.07, *SE* = .64) scored lower than high femininity (*M* = 35.18, *SE* = .41). Neuroticism scores were not significantly different across femininity groups (*F*(1, 365) = 1.54, *p* = .22, *ηp2*= .004).

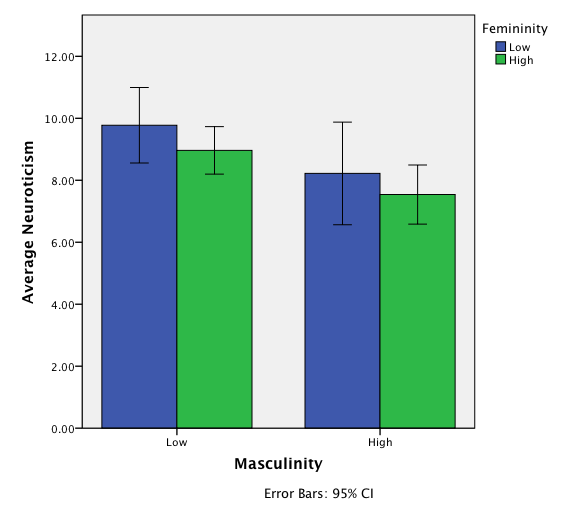
Masculinity scores were significantly different on self esteem (*F*(1, 365) = 73.10, *p* < .001, *ηp2*= .17), and low masculinity participants (*M* = 17.23, *SE* = .25) had higher self-esteem than high masculinity participants (*M* = 13.57, *SE* = .35). Low masculinity (*M =* 35.86, *SE* = .44) participants had higher ratings of the role of women (*F*(1, 365) = 35.18, *p* < .001, *ηp2*= .09), than high masculinity participants (*M* = 31.39, *SE* = .61). Lastly, neuroticism scores were significantly different for different masculinity groups, *F*(1, 365) = 6.12, *p* = .01, *ηp2*= .02. Low masculinity participants (*M* = 9.37, *SE* = .35) scored higher on the neuroticism scale than high masculinity participants (*M* = 7.88, *SE* = .49).



*Figure 1.*

**

*Figure 2.*

**

*Figure 3.*

Example Roy Bargmann Analysis:

Roy Bargmann analyses are used to determine if there is an order to the DVs following up the MANOVA. You are using each step of the analysis to control for variables that you think “occurred first” in time.

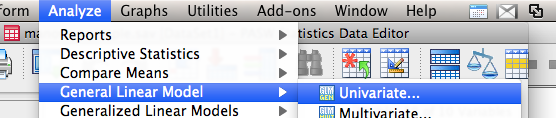
**Roy Bargmann Step Down Procedure**

**Step 1:**The most important DV (DV1) is analyzed as an individual univariate test.

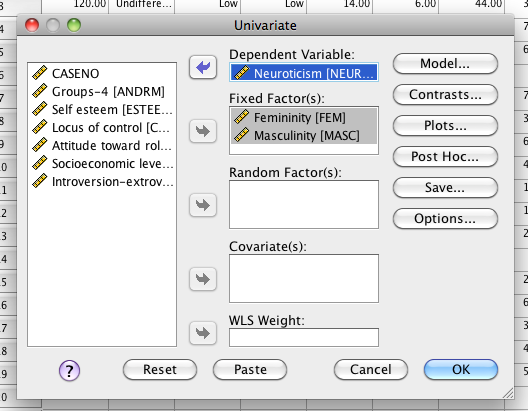
**Step 2:** The second most important DV (DV2), in terms of theoretical importance, is analyzed using the DV1 as a covariate, which controls for the relationship between the two DVs.

**Step 3**: The least important DV (DV3) is assessed using DV1 and DV2 as covariates.

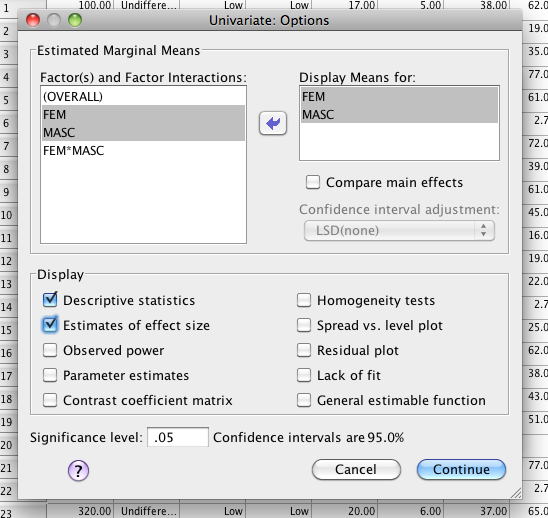
1. Work through the regular MANOVA analysis to find out which main effects and interactions are significant through the multivariate box of the analysis.
2. Take your first DV and run a regular ANOVA on the significant IVs on that DV.
   1. This example we assume that neuroticism is first, and we are going to use FEM and MASC to predict neuroticism (but ignore the interaction because it wasn’t significant at the beginning).
3. Analyze > General Linear Model > Univariate



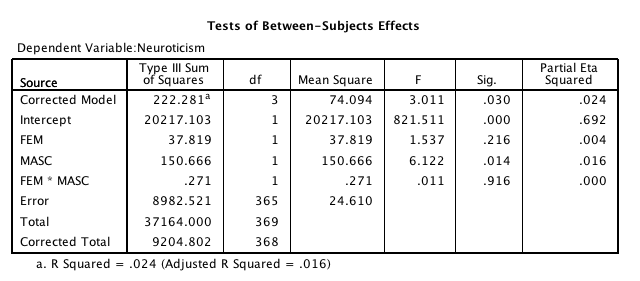
1. Move over your DV into the DV box (neuroticism) and the significant IVs into the fixed factor box.

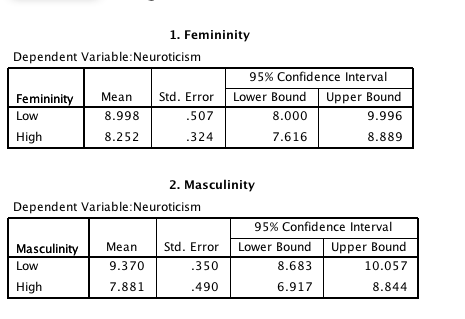


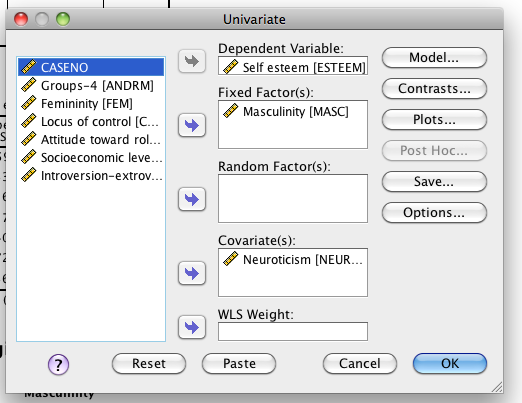
1. Hit Options – ask for means and effect size.

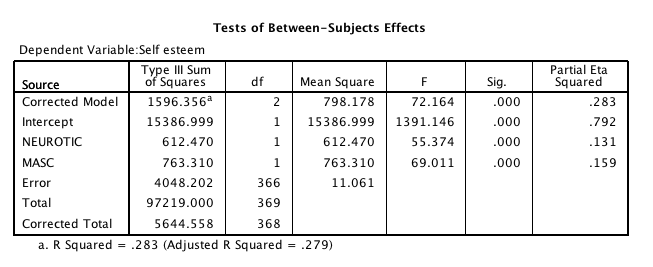


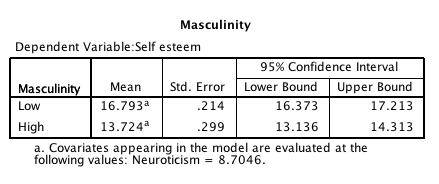
1. Check the Between Subjects box for significant effects.

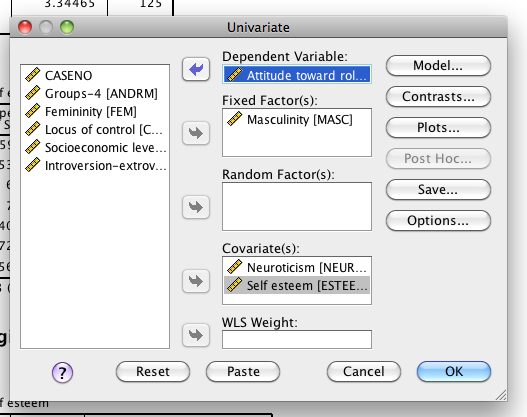


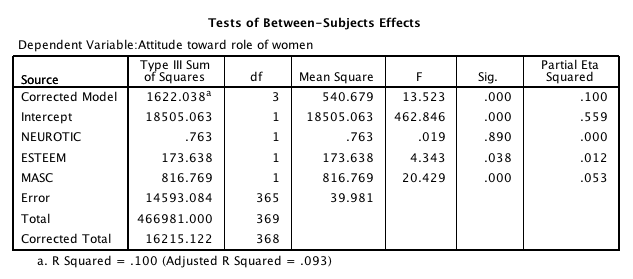


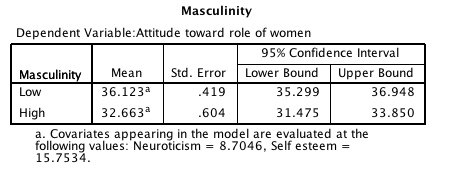












**Effect size and charts are the same as ANCOVA. You would create charts based on the final means in your analysis (i.e. controlling for whatever covariates you ended up with). In the example above, I would use the masculinity LOW versus HIGH with neuroticism and self esteem covaried out – the means would be 36.12, 32.66. Remember you have to create these charts in excel.**

**The write up for a Roy-Bargmann is very similar to the write up above. You would talk about the MANOVA results and then combine with an ANCOVA write up. Here’s an example I found online (note it won’t match our data):**

**Results**

A one way multivariate analysis of variance (MANOVA) performed on three dependent variables. Thus, this design produced one main effect. The omnibus MANOVA test asks whether or not the three dependent variables are associated with the one independent variable. Next, Roy-Bargmann stepdown analysis enables an examination of the pattern of relationships between the dependent (conscientiousness, agreeableness, and neuroticism) variables and the independent variable (gender). As predicted, based on the Wilk’s lambda criterion**,** the combined dependent variables (conscientiousness, agreeableness, and neuroticism) were significantly affected by gender, *F* (3, 172) = 6.74, *p* < .001, *ηp2*= .11.

The Roy-Bargmann stepdown analysis is one technique used to further examine the nature of these relationships. It is analogous to testing the importance of independent variables in multiple regression by sequential analysis. Priorities were assigned to the dependent variables according to theoretical and practical considerations. Dependent variables were ordered as follows: conscientiousness, agreeableness, and neuroticism. This analysis is run by entering the dependent variable with the higher priority first (conscientiousness) in a univariate ANOVA test. After that, the second priority dependent variables are added in an ANCOVA and the first priority dependent variable becomes the covariate. After that, the third priority is entered with the first two covariates.

As predicted, conscientiousness, agreeableness, and neuroticism each made a significant unique contribution to the generalized variance of the dependent variable that best differentiates between gender, *F* (1, 174) = 7.42, *p* < .01, *ηp2* = .04, *F* (1, 173) = 4.38, *p* = .02, *ηp2* = .02, and *F* (1, 172) = 7.80, *p* < .001, *ηp2* = .04, respectively. Specifically, males had higher conscientiousness (conscientiousness *M* = 40.17) than did females (conscientiousness *M* = 37.25). Additionally, males were higher in agreeableness (adjusted *M* = 37.46) than females (adjusted mean = 35.93). Lastly, males neuroticism (adjusted *M* = 14.46) was higher than females (adjusted *M* = 12.10).

Both conscientiousness and agreeablness were also significant in the univariate comparisons (*F* (1, 174) = 7.42, *p* = .01, *ηp2* = .04 and *F* (1, 174) = 11.11, *p* < .001, *ηp2* = .06. Also in the univariate tests, neuroticism was not significant (*F* (1, 174) < 1, *p* = .89, but was significant in the stepdown analysis because the differences in gender that neuroticism accounted for were not also accounted for by the variables entered prior to neuroticism, conscientiousness and agreeableness. Therefore, when these other two variables were treated as covariate, and partialed out, it was found that neuroticism did add a unique contribution. In conclusion, all four of the proposed hypotheses were supported. It was found that overall combination of the personality types, conscientiousness, agreeableness, and neuroticism were affected by gender. Also, significant differences were found according to gender on all three of the personality types.