ES_ANOVA: A SAS® Macro for Computing Point and Interval Estimates of Effect Sizes Associated with Analysis of Variance Models

Jeffrey D. Kromrey, University of South Florida, Tampa, FL
Bethany A. Bell, University of South Carolina, Columbia, SC

ABSTRACT

Measures of effect size are recommended to communicate information on the strength of relationships. Such information supplements the reject/fail-to-reject decision obtained in statistical hypothesis testing. Because sample effect sizes are subject to sampling error, as is any sample statistic, computing confidence intervals for these statistics is a useful strategy to represent the magnitude of uncertainty about the corresponding population effect sizes. This paper provides a SAS macro for computing common effect sizes associated with analysis of variance models. By utilizing data from PROC GLM ODS tables, the macro produces point and interval estimates of eta-squared, partial eta-squared, omega-squared, and partial omega-squared. This paper provides the macro programming language, as well as results from an executed example of the macro.

Keywords: EFFECT SIZES, ANALYSIS OF VARIANCE, CONFIDENCE INTERVALS, BASE SAS®, SAS/STAT®

INTRODUCTION

Long gone are the days when social and behavioral science researchers should simply report obtained test statistics (e.g., t, F, χ^2) and their corresponding *p*-values. Over the years, interpreting importance of scientific research based on the dichotomous reject or fail-to-reject decision has become less popular among some disciplines such as psychology and education. Instead, researchers are encouraged to supplement hypothesis test results with measures of effect magnitude. In fact, according to Thompson (2007), 24 peer-reviewed journals had explicit editorial policies that required authors to include effect sizes or other measures of effect magnitude. A large part of this "cultural change" began over a decade ago when the APA Task Force on Statistical Inference issued their statement that researchers should regularly report effect sizes, calculate confidence intervals, and use graphics to better communicate the nature of their findings for all primary outcomes (Wilkinson & APA Task Force on Statistical Inference, 1999). Unlike *p*-values that are used to determine if an observed effect or relationship is real or due to chance or sampling variability, effect sizes are used to estimate how large the effect or relationship is. Thus, when used together, not only can researchers make statements about the statistical significance of their findings, they can also report on the practical significance of their findings.

More recently, the scholarly discussion around effect sizes has evolved to also include recommendations and formulas for calculating and reporting confidence intervals around effect sizes (Cumming & Finch, 2001; Fidler & Thompson, 2001; Finch & French, 2010; Smithson, 2001; Thompson, 2002). The use of confidence intervals around effect sizes is particularly fruitful for meta-analytic research. For example, just as a confidence interval calculated around a sample mean can generate plausible values for the population mean, a confidence interval calculated around a sample effect size such as Cohen's *d*, *r*, or η^2 can describe plausible values for the population effect size.

Most researchers recognize the value in calculating confidence intervals around effect sizes. However, the difficulty lies in how these confidence intervals must be calculated. Unlike calculating a confidence interval around some parameter estimate such as the sample mean that relies on a standard formula based on either the standard normal distribution or the central *t*-distribution, the calculation of confidence intervals around effect sizes relies on non-central distributions, thus, making the calculations less straightforward for many applied researchers. Fortunately, functions available in the SAS language make calculations using non-central distributions painless.

EFFECT SIZES COMMONLY USED WITH ANALYSIS OF VARIANCE MODELS

Commonly reported effect sizes for ANOVA models include eta-squared $\left(\eta^2\right)$, partial η^2, omega-squared $\left(\omega^2\right)$, and partial ω^2. In general, all four of these effect sizes represent measures of association. However, there are important differences among them. First, both η^2 and partial η^2 are sample effect size estimates, representing the proportion of sample variability in the dependent variable that is associated with variability in an independent variable. These statistics, however, are positively biased as point estimates of the population effect size. The effect sizes ω^2, and partial ω^2, are adjusted to provide better population estimates.

Second, partial and non-partial estimates represent different measures of association. For example, consider a two-way, balanced factorial ANOVA with independent variables of student gender and grade level, and a dependent variable of mathematics achievement. The analysis of variance will provide sums-of-squares (and mean squares) for each of the two main effects (main effects for gender and grade level), the interaction effect, and the within-cell residual or error. *F* tests are formed from ratios of the mean squares to test null hypotheses about main effects and the interaction effect. The total sums-of-squares, representing all of the sample variability in mathematics achievement, is simply the sum of these four components. The formula for η^2 associated with the gender main effect is

$$\eta^2{}_{gender} = \frac{SS_{gender}}{SS_{total}}$$

In contrast, the formula for partial η^2 for this effect is

$$partial\,\eta^2{}_{gender} = \frac{SS_{gender}}{SS_{gender} + SS_{error}}$$

As is evident in the first formula, $\eta^2{}_{gender}$ represents the proportion of the total sample variance in mathematics achievement that is associated with gender. The total sample variance (SS_{total}) includes variability associated with gender, grade level, the interaction between gender and grade level, and the within-cell variability. In contrast, partial $\eta^2{}_{gender}$ represents the proportion of variance in mathematics achievement associated with gender, after the effects of grade level and the interaction between gender and grade level have been removed (note the different denominators for $\eta^2{}_{gender}$ and partial $\eta^2{}_{gender}$). Analogous formulas are used to calculate sample effect sizes for the other two sources of systematic variance in this design – grade level and the interaction between gender and grade level.

Next, to obtain a relatively unbiased estimate of the variance explained in the population by an independent variable, omega-squared can be calculated. Using the same example as above, the formula for ω^2 is presented below.

$$\omega^2_{gender} = \frac{SS_{gender} - (k-1)MS_{error}}{SS_{total} + MS_{error}}$$

where *k* – 1 = the degrees of freedom for the independent variable.

In this formula, ω^2_{gender} estimates the proportion of the population variance in mathematics achievement associated with gender. However, unlike eta-squared, omega-squared takes random error (MS_{error}) into account. Through this process, omega-squared values will be smaller than eta-squared values, with more noticeable differences occurring with smaller samples and/or research designs that include more independent variables.

Similarly, partial omega-squared represents an unbiased estimate of the population proportion of variance in mathematics achievement associated with gender, after the effects of grade level and the interaction between gender and grade level have been removed (note the different denominators for $\omega^2{}_{gender}$ and partial $\omega^2{}_{gender}$).

$$partial\,\omega^2_{gender} = \frac{SS_{gender} - (k-1)MS_{error}}{SS_{gender} + (N-k-1)MS_{error}}$$

CONFIDENCE INTERVALS AROUND COMMONLY USED EFFECT SIZES WITH ANALYSIS OF VARIANCE MODELS
Confidence intervals around simple statistics such as the sample mean can be constructed by adding to and subtracting from the sample mean some number of standard errors. For example, a 95% confidence interval for a mean is estimated as

$$\bar{X} \pm 1.96\sigma_{\bar{X}}$$

The value 1.96 in this formula is obtained from the standard normal distribution and is the number of standard errors that delineates the top and bottom 2.5% of the distribution (hence, the 95% level of confidence).

This approach to confidence interval construction is appropriate for statistics, such as the sample mean, that follow a symmetric sampling distribution. Effect size statistics such as η^2 and ω^2 do not follow a symmetric sampling distribution, necessitating a different approach to confidence interval construction. The interval inversion approach (Steiger, 2004; Steiger & Fouladi, 1997) provides an elegant method that is easily implemented with SAS probability functions. Other approaches to confidence interval construction, such as bootstrap methods, have been proposed. However, empirical investigations of these approaches suggest that they provide few, if any, advantages to the interval inversion approach (see, for example, Finch & French, 2010).

Understanding the interval inversion approach requires a consideration of non-central sampling distributions (i.e., the sampling distributions of t or F when the null hypothesis is false). Figure 1 presents the sampling distributions of the F statistic under four levels of non-centrality (the degrees of freedom for the F statistic were arbitrarily set to 5 and 39 for this illustration). Note that as non-centrality increases, the sampling distribution changes in location, dispersion, and shape. Most importantly, the relative frequencies of the values of F change with changes in non-centrality. As an illustration, a vertical line has been drawn in Figure 1 at the value of 4.46 for obtained F. For the distribution with non-centrality = 0 (i.e., the central F distribution), just less than three-tenths of 1% of the sampling distribution is larger than this value of F. Because this value of non-centrality represents the distribution of F when the null hypothesis is true, obtaining a value of 4.46 is a very rare event if the null hypothesis is true (i.e., when the population means are identical). When the non-centrality value is 5, approximately 6% of the distribution is greater than 4.46; at non-centrality of 10, 19% of the distribution is greater; and at non-centrality of 20, 58% of the distribution is greater. That is, as non-centrality increases this obtained value of F becomes less rare.
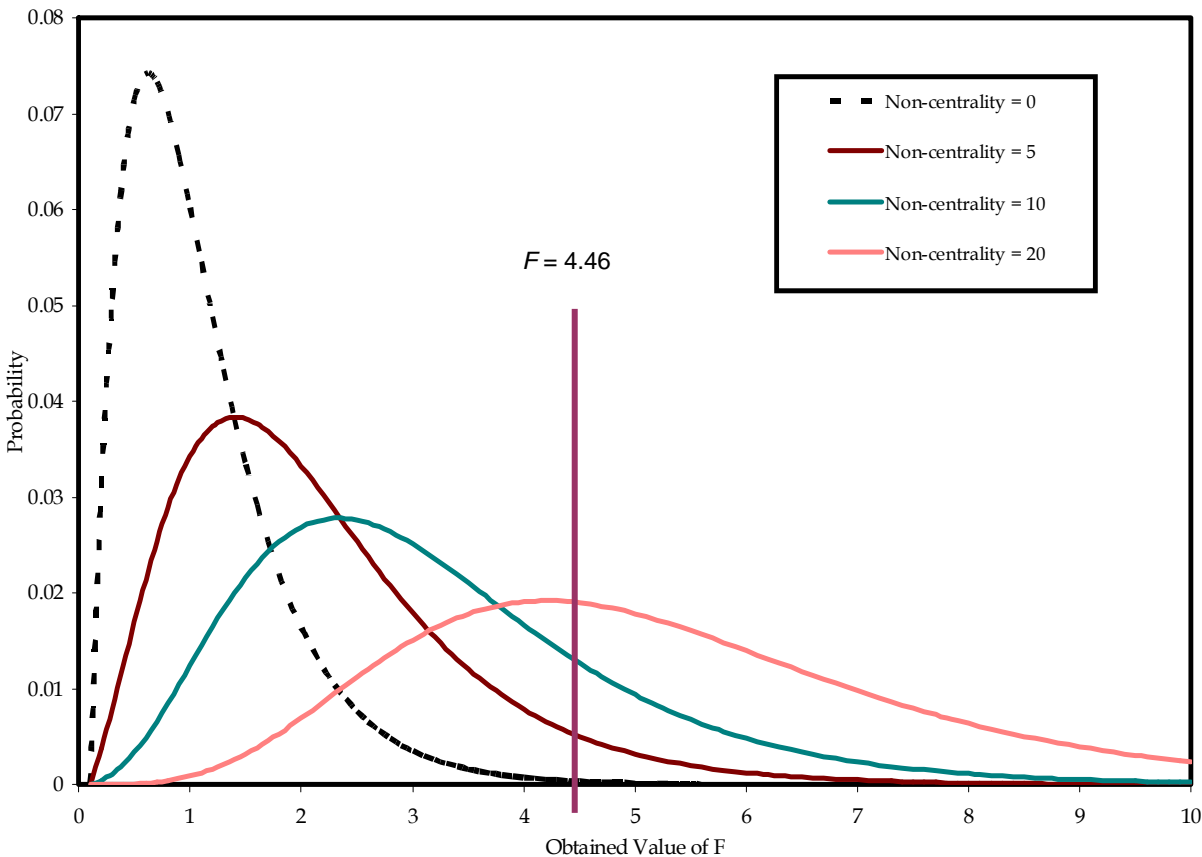

Figure 1. Central and non-central *F* distributions with *df* = 5, 39

These areas under the *F* distribution (or probabilities) can be obtained with the PROBF function in SAS. This function uses four arguments in the following order: the obtained value of *F*, the degrees of freedom for the numerator of *F*, the degrees of freedom for the denominator of *F*, and the non-centrality parameter for the *F* distribution. The last argument is optional if the central *F* distribution (non-centrality = 0) is being evaluated. For example, the statement

<div align="center">Prob = 1 - PROBF(4.46,5,39,20);</div>

provides the area under the curve that is greater than 4.46 when the non-centrality is 20 (that is, Prob = .57748).

Figure 1 provides only four values of non-centrality, but non-centrality parameters have an infinite number of potential values. If we consider a very large number of non-centrality values and compute the proportion of the curve greater than 4.46, we can plot these proportions in reference to the non-centrality values. Such a graph is provided in Figure 2. The horizontal axis provides values of non-centrality and the vertical axis provides the proportion of the *F* distribution that is greater than 4.46 (the probability of obtaining an *F* statistic larger than that obtained, given the non-centrality value).

The interval inversion approach seeks the values of non-centrality that delineate (for example) the top 2.5% and the bottom 2.5% (noted by dotted lines in Figure 2). That is, the value of non-centrality for which 2.5% of the sampling distribution is greater than 4.46 and the value of non-centrality for which 97.5% of the sampling distribution is greater than 4.46 represent the endpoints of a 95% confidence interval around the non-centrality. These non-centrality values are easily transformed into values of η^2 and ω^2.
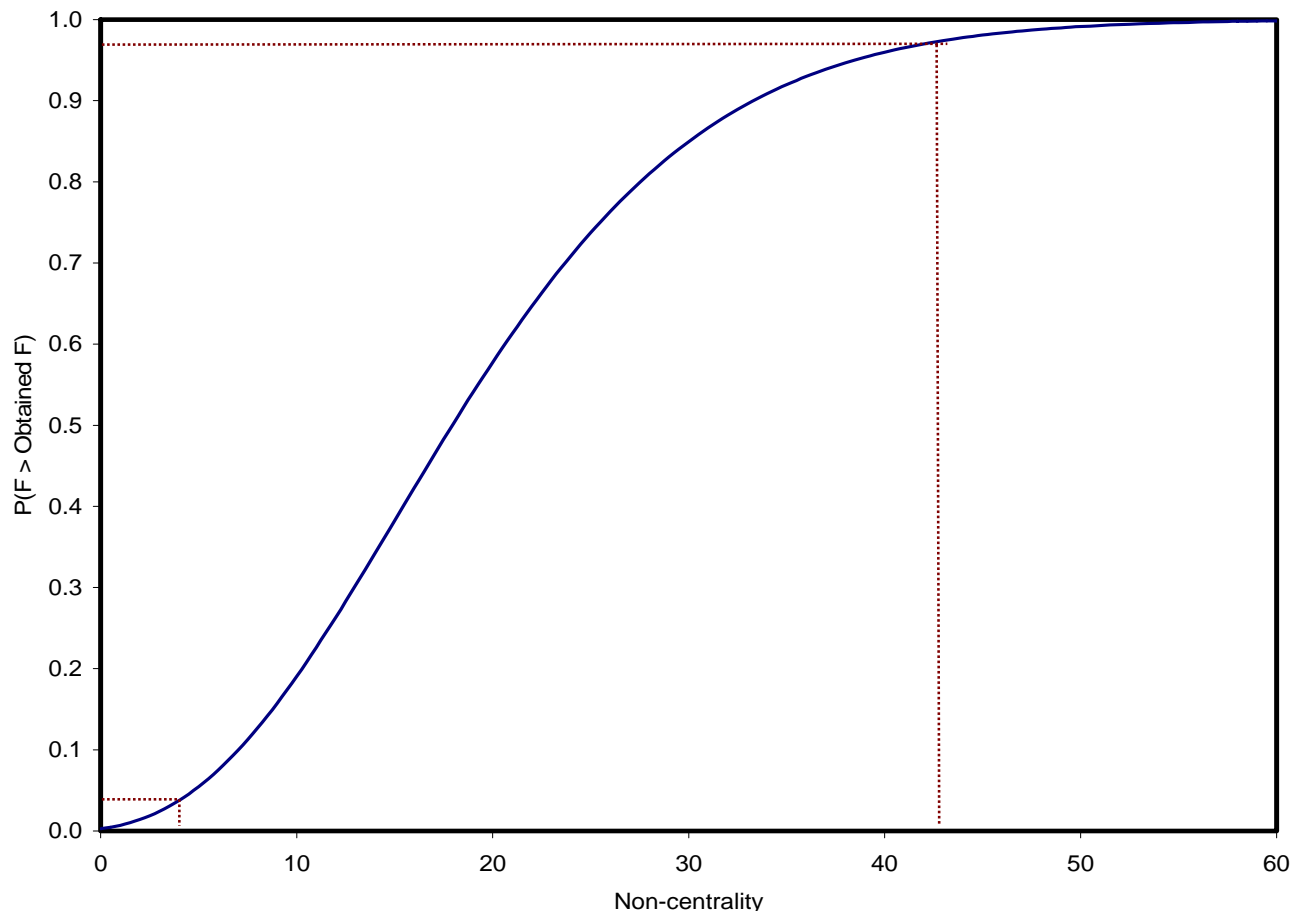

Figure 2. Proportion of the *F* distribution that is greater than 4.46 by non-centrality value

These values of non-centrality (NC) for the *F* distribution can be obtained directly with the FNONCT function in SAS. This function uses four arguments in the following order: the obtained value of *F*, the degrees of freedom for the

numerator of F, the degrees of freedom for the denominator of F, and the desired proportion of the F distribution *smaller* than the obtained value of F. For example, the statement

$$NC = FNONCT(4.46,5,39,.025);$$

provides the non-centrality parameter for which 2.5% of the sampling distribution is less than 4.46 and 97.5% is greater than 4.46 (that is, NC = 43.2386). Similarly, the statement

$$NC = FNONCT(4.46,5,39,.975);$$

provides the non-centrality parameter for which 97.5% of the sampling distribution is less than 4.46 and 2.5% is greater than 4.46 (that is, NC = 3.07142). These values are the endpoints of a 95% confidence interval for the non-centrality parameter.

Converting the endpoints of the interval for non-centrality to the endpoints of the interval for η^2 requires a simple transformation (Cohen, 1988):

$$\eta^2 = \frac{\lambda}{\lambda + df_1 + df_2 + 1}$$

where λ is the non-centrality parameter and df_1 and df_2 are the numerator and denominator degrees of freedom for the F distribution. For this example, the transformation yields endpoints for the η^2 confidence interval of .06389 and .49002, reflecting a relatively large amount of uncertainty about the corresponding parameter.

Although this section has focused on the probability and non-centrality functions for the F distribution (because the F distribution is used in the macro), analogous functions are available in SAS for other sampling distributions. For example, the functions PROBT and PROBCHI provide areas under the curve (i.e., probabilities) for the t distribution and the chi-square distribution, respectively. Similarly, the functions TNONCT and CNONCT provide non-centrality parameters for these distributions. Examples of these non-central distributions applied to interval estimation are provided by Fiddler and Thompson (2001), Reiser (2001), Smithson (2001), Steiger (2004), and Venables (1975).

SOFTWARE LIMITATIONS
Although η^2 and partial-η^2 are automated options available in some statistical software packages, SAS does not provide them in either PROC ANOVA or PROC GLM (with the exception of single factor models). Thus, in order for SAS users to follow the APA Task Force recommendations for reporting effect sizes, extra work is required. For example, using the output from PROC ANOVA or PROC GLM, researchers can calculate these values either by hand or by outputting the ModelANOVA and OverallANOVA ODS tables and calculating the effect sizes through a data step. Thus, in an effort to reduce the burden of calculating commonly reported effect sizes to accompany analysis of variance hypothesis test results, this paper provides a SAS macro to calculate η^2, partial η^2, ω^2, and partial ω^2 for n-way between-subjects ANOVA models. The macro also calculates confidence limits for each effect size and only requires users to have BASE SAS and SAS/STAT.

MACRO ES_ANOVA
ES_ANOVA works with balanced and unbalanced n-way between subjects ANOVA models and contains three arguments: Model, Overall, and Confid. Specifically, the ES_ANOVA user defined macro inputs for the three arguments consist of two ODS table names generated from the PROC GLM procedure (a) ModelANOVA and (b) OverallANOVA, as well as the user's desired level of confidence to be used in calculating the confidence intervals for η^2, partial η^2, ω^2, and partial ω^2. From this information, by default, ES_ANOVA produces two succinct summary tables - one based on Type I sums-of-squares and the other based on Type III sums-of-squares. However, if the user requests Type II and/or Type IV sums-of-squares in PROC GLM, summary tables will also be produced for these sums-of-squares. For each effect included in the ANOVA model, the ES_ANOVA summary tables include the name of the effect along with its corresponding F and p-values, η^2, partial η^2, ω^2, and partial ω^2 effect size estimates, and lower and upper confidence limits for each of the effect size estimates. Because these effect sizes represent proportions of variance, negative values which may occur for lower endpoints of the ω^2 and partial ω^2 confidence intervals are set to zero during the calculation.

The macro ES_ANOVA SAS code is below, followed by an example. ES_ANOVA can also be downloaded from the second author's website (http://www.ed.sc.edu/bell/).

```sas
%macro ES_ANOVA (Model = model_stats, Overall = total_stats, Confid = .95);
filename junk dummy;
proc printto log=junk;
run;
options ls = 148 ps = 47;
OPTIONS FORMCHAR="|----|+|---+=|-/\<>*";
proc format;
   value HT_FMT 1='Type 1' 2 = 'Type 2' 3 = 'Type 3' 4 = 'Type 4';

data total_stats2;
   set &overall (keep=source dependent ss df);
   if source='Corrected Total';
   rename ss=ss_total df = df_total;

data total_error; set &overall (keep=source dependent ss df MS);
   if source='Error';
   rename ss=ss_error df = df_error MS = MS_error;

data model_stats2;
   set &model (keep=ss MS dependent source HypothesisType FValue df ProbF);

proc sort data=total_stats2;
   by dependent;
proc sort data=total_error;
   by dependent;
proc sort data=model_stats2;
   by dependent;

data both;
   merge total_stats2 total_error model_stats2;
   by dependent;
   prob_L = 1-(1-&confid)/2;
   prob_U = (1-&confid)/2;
   conf_pct = 100*&confid;

* +------------------------------------------------------------------+
   Calculating eta-squared and confidence interval around eta-squared
* +------------------------------------------------------------------+;

eta2=ss/ss_total;
dfw_Adj = df_total - df;
MS_E_Adj = (ss_total - SS)/dfw_Adj;
F_Adj = MS/MS_E_Adj;
ncp_eta_L = MAX(0,fnonct(F_Adj,df,dfw_Adj,prob_L));
ncp_eta_U = MAX(0,fnonct(F_Adj,df,dfw_Adj,prob_U));
eta2_L = ncp_eta_L / (ncp_eta_L + df + dfw_Adj + 1);
eta2_U = ncp_eta_U / (ncp_eta_U + df + dfw_Adj + 1);


* +----------------------------------------------------------------------+
   Calculating partial eta-squared and confidence interval around partial
   eta-squared
* +----------------------------------------------------------------------+;

partial_eta2 = ss / (ss + ss_error);
ncp_peta_L = MAX(0,fnonct(FValue,df,df_error,prob_L));
ncp_peta_U = MAX(0,fnonct(FValue,df,df_error,prob_U));
partial_eta2_L = ncp_peta_L / (ncp_peta_L + df + df_error + 1);
partial_eta2_U = ncp_peta_U / (ncp_peta_U + df + df_error + 1);
```

6

```
* +-----------------------------------------------------------------------+
   Calculating omega-squared and confidence interval around omega-squared
* +-----------------------------------------------------------------------+;

Omega2 = (ss - (df*MS_E_Adj)) / (ss_total + MS_E_Adj);
SSB_L=ss_total * eta2_L;
MSW_L=(ss_total - SSB_L) / dfw_Adj;
omega2_L=(SSB_L - (df * MSW_L)) / (ss_total + MSW_L);
    if omega2_L < 0 then omega2_L = 0;
SSB_U=ss_total * eta2_U;
MSW_U=(ss_total - SSB_U) / dfw_Adj;
omega2_U=(SSB_U - (df * MSW_U)) / (ss_total + MSW_U);


* +-----------------------------------------------------------------------+
   Calculating partial omega-squared and confidence interval around partial
   omega-squared
* +-----------------------------------------------------------------------+;

Partial_omega2 = (ss - (df*MS_error)) / (ss_total + MS_error);
SSB_L=ss_total * partial_eta2_L;
MSW_L=(ss_total - SSB_L) / df_error;
P_omega2_L=(SSB_L - (df * MSW_L)) / (ss_total + MSW_L);
if P_omega2_L < 0 then P_omega2_L = 0;
SSB_U=ss_total * partial_eta2_U;
MSW_U=(ss_total - SSB_U) / df_error;
P_omega2_U=(SSB_U - (df * MSW_U)) / (ss_total + MSW_U);

proc sort data = both;
   by dependent HypothesisType;

data both2;
    set both;
    by dependent HypothesisType;
    format HypothesisType HT_FMT.;
    FullLine = repeat('-',144);
    if ProbF >= .0001 then ProbFA = PUT(ProbF,6.4);
    if ProbF < .0001 then ProbFA = '<.0001';


*+-----------------------------------------------------------------------+
         Printed output using FILE PRINT
*+-----------------------------------------------------------------------+;

file print header = H notitles;
if first.dependent|first.HypothesisType then put _page_;
put  @1 Source 20. @22 FValue Best6. @31 ProbFA @40 eta2 Best6.
     @49 eta2_L Best6. @58 eta2_U Best6.
     @67 partial_eta2 Best6. @76 partial_eta2_L Best6.
     @85 partial_eta2_U Best6. @94 omega2 Best6. @103 omega2_L Best6.
     @112 omega2_U Best6. @121 Partial_Omega2 Best6. @130 P_Omega2_L Best6.
     @139 P_Omega2_U Best6.;
return;

H: Put @1 'ANOVA Effect Sizes and Confidence Intervals Produced Using ES_ANOVA
Macro' /
@1 'Dependent Variable:' @30 dependent /
@1 'SS Type:' @30 HypothesisType /
@1 'Confidence Level:' @30 conf_pct 2. @32 '%' //
```

7

```
@1 FullLine /
@46 'Eta-Squared     Partial Eta-Squared           Omega-Squared       Partial
Omega-Sq.' /
@40 '-----------------------   -----------------------   ------------------
-----   -------------------------'/
@1 'Source' @22 'FValue' @31 'Prob>F' @40 'Sample' @49 'Lower' @58 'Upper' @67
'Sample' @76 'Lower' @85 'Upper'@94 'Sample' @103 'Lower' @112 'Upper' @121
'Sample' @130 'Lower' @139 'Upper' /
@1 FullLine;
run;
proc printto log=log;
run;
%mend ES_ANOVA;
```

EXAMPLE OF MACRO ES_ANOVA

Below is an example of PROC GLM code used in conjunction with the macro ES_ANOVA. In this example, the ANOVA model contained one dependent variable (score) and two independent variables (lesson and test). The ANOVA model specified the main effects for lesson and test and the interaction between lesson and test. Thus, the summary tables generated from ES_ANOVA will contain three rows of output – one for each of the effects. The summary tables generated from this example are provided in Figure 3. Separate tables are generated for the Type I and Type III sums-of-squares, which are the default sums-of-squares in GLM (note that the macro provides these tables on two different pages in the output).

```
proc glm;
      ods output ModelANOVA=model_stuff OverallANOVA=total_stuff;
      class Lesson Test;
      model Score = Lesson Test Lesson*Test;
run;

%es_anova (Model=model_stuff, Overall=total_stuff, Confid = .95);
run;
```

In terms of the macro language, in this example, the three user specified inputs for the three macro arguments, Model, Overall, and Confid ,were 'model_stuff', 'total_stuff', and '.95', respectively. These inputs correspond to the names used in the PROC GLM code when creating the ModelANOVA ODS output data file, the OverallANOVA ODS output data file, and our desire to obtain 95% confidence intervals for each of the four effect sizes. The ODS names can be any name that conforms to the SAS syntax requirements and the confidence level can be any value greater than zero and less than 1.00. These three simple elements are the only inputs that a user needs to specify when using ES_ANOVA (note that the macro provides default values of the arguments).

In Figure 3, one can easily review the effect sizes and corresponding confidence intervals, thus, allowing the researcher to make statements about each effect's association with the dependent variable, score. For example, using data from the second table in Figure 3, based on the Type III sums-of-squares, roughly 21% of the total sample variance in the variable score is associated with the main effect of test (η^2_{test} = .2118, 95% CI: 0.0755, 0.3513). When adjusted for sampling variability, the main effect of test accounts for roughly 20% of the variability in the outcome (ω^2_{test} =0.2005, 95% CI: 0.0638, 0.341). Each of the aforementioned values appear in the Type III sums-of-squares summary table, under the headings labeled Eta-Square: Sample, Lower, Upper and Omega-Squared: Sample, Lower, Upper, respectively.

```
ANOVA Effect Sizes and Confidence Intervals Produced Using ES_ANOVA Macro
Dependent Variable:        Score
SS Type:                   Type 1
Confidence Level:          95%


------------------------------------------------------------------------------------------------------------------------
                                      Eta-Squared          Partial Eta-Squared         Omega-Squared          Partial Omega-Sq.
                               -----------------------  -----------------------  -----------------------  -----------------------
Source         FValue  Prob>F  Sample  Lower   Upper    Sample  Lower   Upper    Sample  Lower   Upper    Sample  Lower   Upper
------------------------------------------------------------------------------------------------------------------------

Lesson         10.709  <.0001  0.1346  0.0197  0.2607   0.2112  0.0635  0.3451   0.1126       0  0.2408   0.1213  0.0397  0.3261
Test           34.518  <.0001  0.217   0.0793  0.3565   0.3014  0.143   0.4396   0.2057  0.0676  0.3462   0.2094  0.1309  0.4296
Lesson*Test    11.578  <.0001  0.1455  0.0252  0.2731   0.2245  0.0729  0.3585   0.1237  0.0017  0.2534   0.1321  0.0492  0.3397




ANOVA Effect Sizes and Confidence Intervals Produced Using ES_ANOVA Macro
Dependent Variable:        Score
SS Type:                   Type 3
Confidence Level:          95%


------------------------------------------------------------------------------------------------------------------------
                                      Eta-Squared          Partial Eta-Squared         Omega-Squared          Partial Omega-Sq.
                               -----------------------  -----------------------  -----------------------  -----------------------
Source         FValue  Prob>F  Sample  Lower   Upper    Sample  Lower   Upper    Sample  Lower   Upper    Sample  Lower   Upper
------------------------------------------------------------------------------------------------------------------------

Lesson         11.46   <.0001  0.1441  0.0244  0.2715   0.2227  0.0716  0.3567   0.1222  0.0009  0.2517   0.1307  0.0479  0.3379
Test           33.688  <.0001  0.2118  0.0755  0.3513   0.2963  0.1387  0.435    0.2005  0.0638  0.341    0.2042  0.1265  0.4249
Lesson*Test    11.578  <.0001  0.1455  0.0252  0.2731   0.2245  0.0729  0.3585   0.1237  0.0017  0.2534   0.1321  0.0492  0.3397
```

Figure 3. Type I and Type III sums-of-squares summary tables generated from ES_ANOVA for a 2-way between subjects ANOVA model

CONCLUSION

The macro ES_ANOVA is provided to facilitate the use of effect sizes associated with between-subjects analysis of variance models. Specifically, this macro computes point estimates and confidence intervals for eta-squared, partial eta-squared, omega-squared, and partial omega-squared. Confidence intervals for effect sizes are useful for the interpretation of sample estimates of effect size by providing information about the amount of sampling error expected in the obtained value of the effect size. The macro is easy to use, requiring only the specification of two ODS tables from PROC GLM and the desired level of confidence (by default, 95% confidence intervals are produced). Because the macro is written in Base SAS, it can be modified easily for other applications. For example, graphical displays of both the point and interval estimates of the effect sizes may be added or one-sided confidence intervals may be calculated.

REFERENCES

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cumming G. & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions. *Educational and Psychological Measurement*, *61*, 532-74.

Fidler F. & Thompson B. (2001). Computing correct confidence intervals for ANOVA fixed and random-effects effect sizes. *Educational and Psychological Measurement, 61*, 575-604

Finch, W. H. & French, B. F. (2010, May). *A comparison of methods for estimating confidence intervals for omega squared effect size*. Paper presented at the annual meeting of the American Education Research Association, Denver.

Reiser, B. (2001). Confidence intervals for the Mahalanobis Distance. *Communications in Statistics – Simulation and Computation*, *30*, 37 – 45.

Smithson, M. (2001). Correct confidence intervals for various regression effect sizes and parameters. *Educational and Psychological Measurement, 61,* 605-632.

Steiger, J. H. (2004). Beyond the *F* test: effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods, 9,* 164-182.

Steiger, J. H. & Fouladi, R. T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In Harlow, L. L., Mulaik, S. A. & Steiger, J. H. (Eds.). (1997). *What if there were no significance tests?* Mahwah, NJ: Lawrence Erlbaum, p. 221-257.

Thompson, B. (2002). What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher*, *31*, 25-32.

Thompson, B. (2007). Effect sizes, confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools, 44*, 423-432.

Venables, W. (1975). Calculation of confidence intervals for noncentrality parameters. *Journal of the Royal Statistical Society*, *Series B, 37*, 406-412.

Wilkinson & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594-604.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the first author at:

Jeffrey D. Kromrey
University of South Florida
4202 East Fowler Ave., EDU 105
Tampa, FL 33620

Work Phone: (813) 974-5739
Fax: (813) 974-4495
Email: kromrey@usf.edu