Data Size Planning for Multifactor ANOVA Designs via Adequately Narrow Confidence Intervals for Partial Eta-Squared

Reza Norouzian
University of Texas at Austin

Citation for the following *AERA* paper:

Norouzian, R., de Miranda, M. A. (2019). Data size planning for multifactor ANOVA designs via adequately narrow confidence intervals for partial eta-squared. Paper accepted at the annual meeting of the *American Educational Research Association* (***AERA, 2019***), Toronto, Canada.

Data Size Planning for Multifactor ANOVA Designs via Adequately Narrow Confidence Intervals for Partial Eta-Squared

Reza Norouzian

Michael de Miranda

Perspective

Researchers traditionally plan for the size of their required sample size to ensure a sufficient level of statistical power is achieved (Cohen, 1988). A goal of this power-analytic approach (PAA) is to confidently distinguish statistically significant effects from non-significant effects. Through time, PAA has fallen short of accommodating one of the main goals of educational researchers (Cumming & Calin-Jageman, 2017). Specifically, while researchers often aim to accurately estimate the actual size of an effect of interest "an accurate estimate need not be significant and a significant estimate need not be accurate" (Kelley & Rausch, 2006, p. 380). Methods of data size planning via accuracy in effect size estimation (AESE) focus on ensuring that researchers obtain adequately *narrow* confidence intervals for their measures of effect size.

Significance of the Study

Analysis of variance (ANOVA) is a common statistical tool in educational research (Troncoso Skidmore & Thompson, 2010). However, AESE methods of data size planning for multifactor ANOVA designs using common effect size measures have not been well developed. Given (a) the widespread use of partial eta-squared, as a measure of effect size in the context of multifactor ANOVA designs by researchers (Norouzian, 2015; Norouzian, De Miranda, & Plonsky, 2018; Norouzian & Eslami, 2013a, 2013b, 2016; Norouzian & Plonsky, 2018a; Norouzian & Plonsky, 2018b), (b) the advances made in the area of statistical computing to achieve fast implementation of AESE methods, and (c) the cost of recruiting human participants for quantitative research purposes, the proposed methods in this article enable researchers to fulfill the goal of gaining accuracy in estimating a critical effect size measure in their intended, multifactor studies.

Objectives

To advance data size planning in multifactor ANOVA designs, we take three steps. First, we provide a method that determines the minimum required data size so that the *expected* width of the confidence interval for partial eta-squared (H^2) effect size in multifactor ANOVA designs will be adequately narrow. Then, a second method is developed that determines the minimum required data size so that the *observed* width of the confidence interval for partial eta-squared (H^2) effect size in the target designs will be adequately narrow given a preset level of probabilistic certainty. Finally, we provide an R program that enables the practical use of the methods discussed in this article.

Methods

The estimator of H^2, η_p^2, is known to be upwardly biased (Grissom & Kim, 2012). Therefore, η_p^2 can overestimate the population omnibus effect of a factor in a multifactor design given *K* levels for that factor, *G* total number of groups in the design, and *N* total sample size.

Due to bias, H^2, as a population parameter, and the expected value of its estimator, η_p^2, do not represent the same value. To obtain the exact expected value of η_p^2 via definition of expectation of continuous random variables as

$$E\left[\eta_p^2|H^2,G,N,K\right]=\int_0^1 \eta_p^2\times f\left(\eta_p^2 \mid H^2,G,N,K\right)d\eta_p^2 \tag{1}$$

we will need to have *f*(**.**), the probability density function (pdf) of η_p^2. One may use a jacobian transformation to compute the pdf of η_p^2 from the widely known pdf of the *F* statistic and then use equation 1 to obtain the exact expected value of η_p^2. The required steps appear in the appendix.

For convenience, we use $E[\eta_p^2]$ to denote $E\left[\eta_p^2|H^2,G,N,K\right]$ throughout the present article. As is discussed in the next section, removing the bias in η_p^2 is critical in planning for the data size such that the *expected* width of the confidence interval for H^2 is ensured. A closed-form derivation for confidence interval for H^2 effect size is not available. However, a reliable solution can be found using an optimization routine (Reza Norouzian & Luke Plonsky, 2018). As expected, the width of the confidence interval for H^2, *w*, is obtained by subtracting the lower limit (H_L^2) from the upper limit (H_U^2):

$$w=H_U^2-\ H_L^2 \tag{2}$$

However, the relationship between η_p^2 and *w* is non-monotonic. For example, in a two-factor ANOVA design where K = 4, N = 120, G = 12, Figure 1 displays the width of the 95% confidence interval for H^2 as a function of η_p^2 in the design.

Figure 1 Here

Figure 1 illustrates, as values of η_p^2 increase up to .3392, the width of the confidence interval consistently increases. However, as values of η_p^2 increase beyond .3392, the width of the confidence interval continuously decreases, albeit at a lower rate. Indeed, for other reasonable combinations of design elements, *K*, *G*, and *N*, a similar relationship between η_p^2 and the width of the confidence interval mostly peaked between .33 and .43 is observed. This non-monotonic relationship indicates that keeping *K, G,* and *N* constant, for the H^2 value at the peak, the minimum required sample size to achieve a certain width will be the largest. Understanding the relation between η_p^2 and the width of the confidence interval for H^2 will prove crucial in devising a method for determining the minimum required data size for a desired, *observed* width with a preset level of probabilistic certainty.

Planning data size for narrow, expected confidence interval around H^2

The expected confidence interval width for H^2 is defined as the width of the interval obtained when using $E[\eta_p^2]$ for the observed η_p^2. That is

$$E\left[H_U^2|E\left[\eta_p^2\right]\right] - E\left[H_L^2|E\left[\eta_p^2\right]\right] = E\left[H_U^2 - H_L^2|E\left[\eta_p^2\right]\right]. \tag{3}$$

Our goal is to plan for the total sample size *N* such that the expected width of the interval, $E\left[H_U^2 - H_L^2|E\left[\eta_p^2\right]\right]$ (hereafter denoted $E[w]$), will be no wider than some desired width (*W*). If η_p^2 was an unbiased estimator of H^2, then $E[w]$ could have been obtained by using a researcher supplied H^2 for $E[\eta_p^2]$. Because without bias, H^2 and $E[\eta_p^2]$ would have been the same value. But in the presence of bias, $E[\eta_p^2]$ must be obtained by applying equation 1 to H^2 in the data size planning method discussed next.

Since no closed-form derivation of the confidence interval for H^2 is known, the data size planning for achieving the desired confidence interval width, *W*, for H^2 in multifactor ANOVA designs depends on an advanced optimization procedure. The goal is to find the minimum required total sample size such that $E[w] \leq W$. Below, we provide a conceptual summary of the optimization approach. The optimization approach requires converting the researcher provided H^2 to an *F* variate of a noncentral *F* distribution while making use of the algebraic relations among the total sample size (*N*), non-centrality parameter (*ncp*), and the error degree of freedom (*df2*) to solve for *df*2. Then, *df2* is solved for such that the desired width of the interval, *W*, requested by the researcher is the best numerically possible interval for the current design elements. Next, the obtained *df2* is converted to *N* realizing that in multifactor ANOVA designs $N = df2 + G$. From this step, an initial *N* is obtained. Lastly, the process is repeated using $E[\eta_p^2]$ corrected via equation 1 in place of the researcher supplied H^2 updating the initial *N* produced in the previous step.

Planning Data Size for Narrow, Observed Confidence Interval around H^2

The emphasis on the *expected* width, $E[w]$, in the previous section should reveal that the sample size obtained via the previous method only ensures obtaining a width for the confidence interval for H^2 that will be no wider than the average width of the intervals obtained across infinitely many replications. Thus, no guarantee that in any one case, the *observed* confidence interval will have the width desired by the researchers. Below, a second method is developed through which obtaining a width for the confidence interval for H^2 no wider than desired is ensured in any one case with a specified level of probabilistic certainty.

Our goal is to be certain with θ probability that no matter what η_p^2 value for the factor of interest is observed by the researcher, the width of the observed confidence interval, w, for H^2 will be equal to or smaller than the desired width, W. That is,

$$p(w \leq W) \geq \theta. \tag{4}$$

Stated differently, our method should be devised such that θ% of the time the observed width of the confidence interval for H^2 is no wider than W, while there will be (1 – θ)% of the time in which this is not the case. One obstacle of achieving such a method is the non-monotonic relationship between η_p^2 and the width of the confidence interval for H^2 (see above). Let us suppose that we planned for the data size using the method of expectation. As shown in Figure 2, given $E[\eta_p^2] = .6$, $N_{(planned)} = 120$, $K = 4$, and $G = 12$, if in the actual study, the researcher observes an η_p^2 falling between .1334 and .6, then $w > W$, an undesirable case. However, if the researcher observes an η_p^2 value larger than .6 or smaller than .1334, then $w < W$, both desired cases. Thus, for any set of elements (N, K, G, and $E[\eta_p^2]$), there will be η_p^2 values that if observed, will result in an undesired situation, $w > W$.

Figure 2 Here

What follows are the details of a second method that determines the data size such that $w \leq W$ in any case with a specified level of probabilistic certainty is ensured.

Given the N obtained for the expected width of the confidence interval for H^2 along with K, G, and $E[\eta_p^2]$, we can find the upper and the lower limits of a (2θ – 1)% confidence interval for H^2. For any given set of N, K, G, and $E[\eta_p^2]$, the limits of such a two-sided confidence interval will reveal the η_p^2 values that have (1 − θ)% probability of occurring on either side of the sampling distribution of η_p^2. If these limit values, denoted H_L^{2*} and H_U^{2*}, take the place of $E[\eta_p^2]$

used in the method of expectation, then two new data sizes, denoted N_{L*} and N_{U*}, are obtained. The larger of N_{L*} and N_{U*} will be the minimum necessary sample size that ensures obtaining an observed width, w, no wider than the desired width, W, with θ probability, *if and only if* the observed η_p^2 for the omnibus effect falls outside the $H_L^2{}_*$ and $H_U^2{}_*$ limits. To remove this condition, it is critical to remember (see Figure 1) that for any set of design elements, the largest possible N is obtained for the value of η_p^2 that maximizes the non-monotonic relation between width of the confidence interval for H^2 and η_p^2, denoted $\eta_p^2{}_{max}$. If after determination of $\eta_p^2{}_{max}$, it is found that $\eta^2_{p\ max}$ does not fall within $\left[H^2_{L*}, H^2_{U*}\right]$, then the larger of N_{L*} and N_{U*} will ensure observing a confidence interval width, w, no wider than the desired width, W, with θ probability. However, if after determination of $\eta_p^2{}_{max}$, it is found that $\eta^2_{p\ max}$ does fall within $\left[H^2_{L*}, H^2_{U*}\right]$, then limits of $(1 - \theta)$% two-sided confidence intervals will also need be obtained to ensure that $p(w \leq W) \geq \theta$. If these limit values, denoted $H_L^2{}_{**}$ and $H_U^2{}_{**}$, take the place of $E[\eta_p^2]$ used in the method of expectation, then two yet newer data sizes, denoted N_{L**} and N_{U**}, will be obtained. The largest of N_{L*}, N_{U*}, N_{L**}, and N_{U**} will represent the minimum sample size that ensures observing an x% confidence interval for H^2 respecting $w \leq W$ with θ probability.

Methods Implementation

A comprehensive R function named `plan.f.ci()` has been developed to facilitate the implementation of the methods of data size planning discussed in the present article. The function (hosted by an anonymous online account) can be sourced into R via:

```
source("https://raw.githubusercontent.com/izeh/i/master/n.r")
```

For example, suppose an educational researcher believes that H^2 (`H2`), the population omnibus effect of a teaching approach with 4 levels (`n.level`), in an intended 2 × 4 ANOVA design

(`design`) is 0.2. To plan for the number of required participants such that with 99% certainty (`assure`), s/he obtains a 95% confidence interval (`conf.level`) for H^2 that is no wider than 0.15 (`width`), the R function `plan.f.ci()` may be used as

```
plan.f.ci(H2 = .2, width = .15, n.level = 4, conf.level = .95, design = 2 * 4, assure = .99)
```

All assumptions being met, it turns out that if our researcher would be able to recruit 344 participants, then the goal of accuracy in effect size estimation in his/her study is guaranteed to be achieved with 99% certainty.

Conclusion

Methods developed in the current article directly help researchers in planning to achieve their critical goal of accurately estimating the size of their effects of interest regardless of statistical significance of those effects. We hope that the usefulness of these methods helps promoting the application of AESE in planning for various forms of educational research.

References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cumming, G., & Calin-Jageman, R. (2017). *Introduction to the new statistics: estimation, open science, and beyond*. Chicago: Routledge.

Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications* (2nd ed.). NY, New York: Routledge.

Norouzian, R. (2015). Does teaching experience affect type, amount, and precision of the written corrective feedback? *Journal of Advances in English Language Teaching, 3*(5), 93-105.

Norouzian, R., De Miranda, M., & Plonsky, L. (2018). The Bayesian revolution in second language research: An applied approach. *Language Learning, 68,* 1032 – 1075.

Norouzian, R., & Eslami, Z. (2013a). Does synchronous computer-mediated communication improve EFL learners' oral proficiency? *Modern English Teacher, 22*(3), 40-42.

Norouzian, R., Eslami, Z., R. (2013b). Applying teacher feedback: Grounded theory perspective. *Tennessee Foreign Language Teaching Association Journal, 4,* 72 – 87.

Norouzian, R., & Eslami, Z. (2016). Critical perspectives on interlanguage pragmatic development: An agenda for research. *Issues in Applied Linguistics, 20*(1), 25-50.

Norouzian, R., Mehdizadeh, M. (2013). Reading strategy repertoires in EAP contexts: students and teachers in academic reading strategy use. *International Journal of Language Learning and Applied Linguistics, 3,* 5 – 12.

Norouzian, R., & Plonsky, L. (2018a). Correlation and simple linear regression in applied linguistics. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave* *handbook of applied linguistics research methodology* (pp. 395-421). UK, London: Palgrave.

Norouzian, R., & Plonsky, L. (2018b). Eta- and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research*. doi:10.1177/0267658316684904

Skidmore, S., & Thompson, B. (2010). Statistical Techniques Used in Published Articles: A Historical Review of Reviews. *Educational and Psychological Measurement, 70*(5), 777-795. doi:10.1177/0013164410379320

Appendix

The computation of $E[\eta_p^2]$ by first obtaining the pdf of η_p^2 random variable from the widely known pdf of *F* random variable can be done in three steps realizing throughout that *df1* = *K* – 1, and *df2* = *N* – *G*.

Step 1, define the non-centrality parameter (*ncp*) of the pdf, *f*(.), of the *F* distribution given H^2:

$$ncp = \frac{H^2 N}{1 - H^2} \tag{5}$$

Step 2, form the jacobian of the transform and scale the pdf of the non-central *F* distribution:

$$d = \frac{df2}{df1} \tag{6}$$

$$p = d\left(\frac{\eta_p^2}{1 - \eta_p^2}\right) \tag{7}$$

$$f\left(\eta_p^2 \mid H^2, df1, df2\right) = f(p \mid ncp, df1, df2) \times d \times \left[\left(\frac{1}{1 - \eta_p^2}\right) + \left(\frac{\eta_p^2}{\left(1 - \eta_p^2\right)^2}\right)\right] \tag{8}$$

Step 3, get the $E[\eta_p^2]$ via the definition of expectation of continuous random variables realizing that *K* = *df1* + 1, and *N* = *df2* + *G*:

$$E\left[\eta_p^2\right] = \int_0^1 \eta_p^2 \times f\left(\eta_p^2 \mid H^2, G, N, K\right) d\eta_p^2 \tag{9}$$

Figures to be Inserted in Text

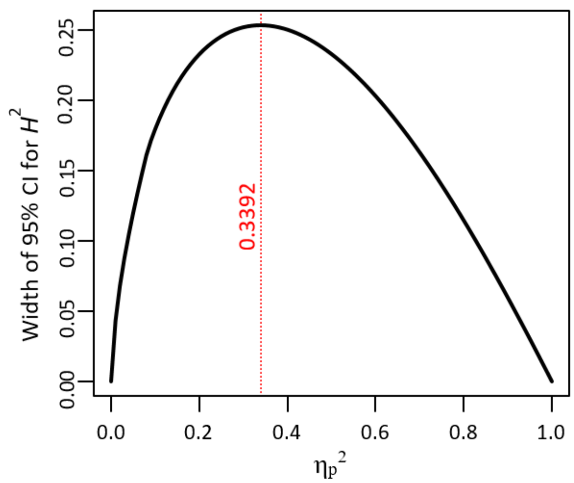

Figure 1. Non-monotonic relationship between the 95% confidence interval (CI) width and η_p^2 given K = 4, N = 120, G = 12.

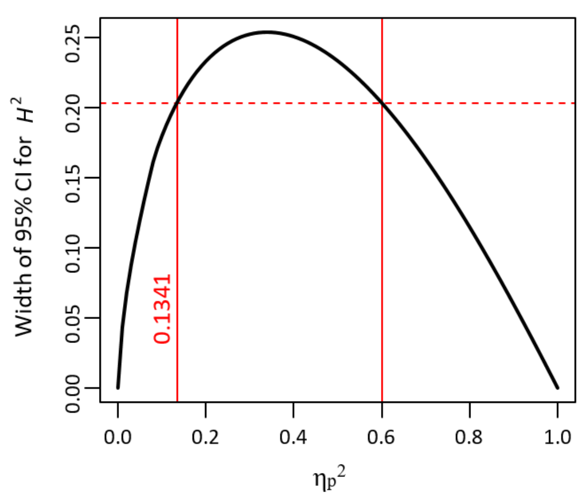

Figure 2. Regions of desired and undesired width in the relationship between the 95% confidence interval (CI) width and η_p^2 given K = 4, G = 12, and N = 120.