Statistical Power in the *Journal of Agricultural Education*, 2012 – 2022

Henry Akwah
Graduate Assistant
University of Arkansas
1120 W. Maple St.
AFLS E111
Fayetteville, AR 72701
(479) 575-2035
hoakwah@uark.edu

Christopher M. Estepp
Associate Professor
University of Arkansas
1120 W. Maple St.
AFLS E111
Fayetteville, AR 72701
(479) 575-2035
estepp@uark.edu

Donald M. Johnson
University Professor
University of Arkansas
1120 W. Maple St.
AFLS E111
Fayetteville, AR 72701
(479) 575-2035
dmjohnso@uark.edu

Statistical Power in the *Journal of Agricultural Education*, 2012 - 2022

Statistical power has been defined as the probability of finding a statistically significant result with an inferential statistical test when an effect actually exists in the population (Turner & Houle, 2018). Stated differently, statistical power is the probability of avoiding a Type II error, which occurs when a researcher fails to reject the null hypothesis when the null hypothesis is false in the population (Glass & Hopkins, 1996). Statistical power increases as sample size gets larger, the alpha level of the statistical test increases (i.e., from .05 to .10), and when an amplified magnitude of the effect occurs in the population (Glass & Hopkins, 1996). While the latter method of increasing power is outside the direct control of the researcher, the first two can be controlled. Cohen (1988) recommended a minimum power of .80 when conducting inferential statistical tests. However, because increasing the alpha level directly increases the probability of committing a Type I error (rejecting a true null hypothesis), the preferred method of increasing statistical power is to increase sample size (Turner & Houle, 2018).

Theoretical Framework

This research was framed using Ajzen's (1991) Theory of Planned Behavior (TPB). According to the TPD, an individual's decision to engage in a specific behavior depends on their attitude, their subjective norms, and their perceived behavioral control related to the behavior. This research aims to increase the profession's positive attitudes, subjective norms, and perceived behavioral control toward the intended behavior of increased consideration of statistical power when planning and conducting research employing inferential statistics.

Purpose

The purpose of this study was to determine the statistical power (at the small, medium, and large effect sizes) for inferential statistical tests reported in articles published in the *Journal of Agricultural Education* (*JAE*), 2012 – 2022.

Methods

The researchers manually examined each published article in *JAE* between 2012 and 2022, inclusive, and identified all articles where inferential statistics were used. A coding sheet was developed, and the following data was collected for each inferential statistical test reported; the specific statistical tests used, the total number of subjects included in each analysis, the number of subjects per group, and the stated alpha level. For multiple regression, the number of predictor variables in the model were also recorded, and for MANOVA, the number of dependent variables was also recorded. G*Power Version 3.1.9.2 (Faul et al., 2007) software was used to calculate the statistical power for each inferential statistical test at the small, medium, and large effect sizes as summarized by Kotrlik et al. (2011).

Results

Seventy inferential statistical tests (See table 1) were reported in *JAE* for the 11 years between 2012 and 2022. The most frequently reported tests were bivariate correlations (27.1%), independent *t*-tests (24.6%), and one-way ANOVAs (22.9%), while one-way MANOVAs

(2.9%), and bivariate regression (1.4%) were the least commonly reported inferential statistics. At the small effect size, mean statistical power ranged from 0.101 (one-way MANOVA) to 0.625 (paired *t*-test), while at the medium effect size, mean statistical power ranged from 0.310 for factorial ANOVA to 0.989 for paired *t*-tests. Lastly, at the large effect size, mean statistical power ranged from 0.680 (factorial ANOVA) to 1.0 for both paired *t*-tests and bivariate regression.

Table 1

Descriptive Statistics for Mean Power by Test and Effect Size

		Effect size					
		Small		Medium		Large	
Statistical Test	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Independent *t*-test	17	0.314	0.300	0.667	0.332	0.867	0.214
Paired *t*-test	4	0.625	0.305	0.989	0.019	1.000	0.000
One-way ANOVA	16	0.403	0.331	0.812	0.220	0.968	0.068
Factorial ANOVA	3	0.086	0.007	0.310	0.024	0.680	0.012
One-way MANOVA	2	0.101	0.019	0.436	0.095	0.855	0.068
Factorial MANOVA	4	0.428	0.438	0.650	0.433	0.778	0.374
Multiple regression	4	0.415	0.388	0.931	0.076	0.999	0.001
Bivariate correlation	19	0.333	0.291	0.861	0.192	0.986	0.039
Bivariate regression	1	0.346	-	0.990	-	1.000	-

Conclusions and Recommendations

For small effects none of the tests reached the minimum recommended statistical power of 0.80 (Cohen, 1988). The mean power for independent *t*-tests, factorial ANOVA, one-way MANOVA, and Factorial MANOVA were less than the recommended statistical power at the medium effect size; however, paired t-tests, one-way ANOVA, multiple regression, and bivariate regression all had power greater than .80. At the large effect size, factorial ANOVA and factorial MANOVA were the only statistical test reported with a mean power less than Cohen's recommendation. Overall, tests reported in *JAE* between 2012 and 2022 were capable of detecting large effects at an acceptable level, however, were less effective for medium and small effects. Researchers should consider statistical power as studies are planned and select appropriate sample sizes to ensure acceptable power for the anticipated effect size. Additionally, researchers should report statistical power for all inferential analyses. This would enable other researchers to determine if reported non-significant differences were due to no effect in the population or low statistical power.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

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

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Glass, G.V., & Hopkins, K.D. (1996). *Statistical methods in education and psychology* (3rd ed.). Allyn & Bacon.

Kotrlik, J.W., Williams, H.A, & Jabor, M.K. (2011). Reporting and interpreting effect size in quantitative agricultural education research. *Journal of Agricultural Education, 52(1), 132-142. https://doi.org/10.5032/jae.2011.01132*

Turner, D., & Houle, T.T. (2018). The importance of statistical power calculations. *The Journal of Head and Face Pain, 58*(8), 1187-1191.