

Correlations: Description or Inference?

Larry E. Miller, Professor
The Ohio State University

This paper was solicited and the review and publication in The Journal managed by the Editor .

Research is often conducted in agricultural education with the purpose of explaining and/or predicting phenomena. This research is frequently labeled associational, relational, correlational, ex post facto or causal-comparative depending primarily upon which author one studied when learning about research. The purpose of explaining is most often addressed by what is known as the correlation question: "Do two variables vary together?" The second, a prediction question, examines: "How well can 'Y' be predicted from knowledge of 'X'?" As might be expected, the statistic used to answer the correlation question is a correlation and regression analysis is used to address the second question. These types of research are called descriptive by most research writers. No cause-and-effect (causality) relationships can be drawn since the independent variables (causes) are not under the control of the researcher, but are naturally occurring or self-selected by the subjects. The researcher does not "control" which subject gets which level of the independent variable or variables. Since the researcher cannot exercise control, the research is descriptive. Further, the end sought (purpose) of the research is description in contrast to explanation or control of outcomes.

The Problem

This article will examine relational research conducted to describe or to explain phenomena. In other words, the research is attempting to examine how things vary together. In the examples, the variable predicted from, the independent variable, is labeled 'X'; and the variable predicted to, the dependent variable, is labeled 'Y'. Further, the Pearson product moment correlation coefficient (r) is used as the example case.

A problem seems to exist within agricultural education research about the descriptive versus the inferential application of the correlation statistic used to examine the correlation question. An analysis of articles from Volume 31 to present of the Journal of Agricultural Education revealed

30 which reported correlations. Of these 30, 18 articles (60%) reported significance (probability = p) levels associated with the correlations. In no case among the 18 articles was the research or null hypothesis stated related to the reported test. Was the intent of the writers of these articles to describe only or also to test hypotheses (make inferences)? Since no hypotheses were stated, one must assume that they sought to use the correlation descriptively. When researchers report a significance level associated with r , they are testing, whether they know it or not, that $H_0: p = 0$. If researchers wish to make inferences to populations based upon data from a sample, then they can test hypotheses regarding the correlations in the population. Such hypotheses are rarely stated, however. This problem is further exacerbated when negligible or low correlations are later discussed as "significant." When r is reported, it is describing the direction and magnitude of the relationship. Similar problems have been observed in papers presented at national and regional research meetings in agricultural education.

Discussion and Conclusions

Let us begin by looking at the use of correlations for description using r as the example. First of all, the assumptions underlying the use of r should always be addressed. Researchers should always construct a scatterplot before calculating r to investigate whether the relationship is linear, both X and Y are at least on the interval scale of measurement, the two variables have similar shapes of distributions, there is sufficient variance in both variables to demonstrate relationship if the two variables are correlated, and the data are homoscedastic. If the data are curvilinear or heteroscedastic, then r should not be calculated or reported. Writers should confirm that these assumptions have been met. Secondly, researchers should indicate the direction and magnitude of the correlations which were calculated. A standard convention should be adopted for describing the magnitude of the correlations. An example is provided in Table 1 (Davis, 1971).

A positive correlation coefficient indicates that the line of best fit going through the scatterplot

Table 1. Adjectives for describing the magnitude of correlations

<i>r</i>	Adjectives
1.0	Perfect
.70 to .99	Very High
.50 to .69	Substantial
.30 to .49	Moderate
.10 to .29	Low
.01 to .09	Negligible

goes from lower left to upper right, which indicates that as X increases then Y increases. A negative correlation coefficient indicates that the line of best fit going through the scatterplot goes from upper left to lower right, which indicates that as X increases then Y decreases. Thus, the Pearson *r* indicates the degree and direction of linear relationship between X and Y.

Writers could provide further description by calculating the coefficient of determination (r^2) and/or the coefficient of nondetermination ($1 - r^2$). The coefficient of determination is a simple, good and easily remembered method of describing the magnitude of the correlation. In essence, r^2 describes the amount of variability in Y which is explained by knowledge of X. In other words, if one knows X, then how nearly can Y be predicted? For example, if $r = .9$, the $(.9)^2 = .81$; which can be multiplied times 100 and interpreted like a percentage, i. e., 81% of the variability in Y has been explained. How much has not been explained? Obviously, to answer this question, the researcher would calculate the coefficient of nondetermination: $(1 - r^2) = (1 - .9^2) = (1 - .81) = .19$; $.19 \times 100 = 19\%$ of the variability in Y is unexplained by knowledge of X.

Correlations of the magnitude of .9 are not very common in research in agricultural education. However, a .20 might be more frequent. The coefficient of determination for $r = .2$ would be $r^2 = .04$; $.04 \times 100 = 4\%$. This would be interpreted that when the $r = .2$, then the amount of variability in Y accounted for by knowledge of X is only 4%, which leaves 96% not explained. In short, a correlation coefficient of .2 is much more common in agricultural education research, but X is not practically important in explaining the variability in

Y. Knowledge of X is of little help in explaining the variability in Y. A correlation between Supervised Agricultural Experience (SAE) Quality and FFA achievement of .2 would, in essence, say that knowledge of SAE quality of a student would not help explain the variability in how that student might achieve in FFA.

Note, however, that if enough students were studied the .2 correlation could be called "significant" by a researcher. Ferguson (1976) presents a table of critical values of the correlation coefficient, alpha = .05, 1-tailed test, which indicates that if a researcher studied 102 students (100 df), then a correlation coefficient of .195 would be of sufficient magnitude to be called significant. The .195 (critical *r*) value is less than .2 (observed *r*) and is illustrative of the earlier example where the practical significance of the value was shown to be minimal in explaining the variability in Y from knowledge of X.

Perhaps some researchers originally set out to describe the relationships among variables and to be able to predict who might achieve well in the FFA with a sample of 300 students. However, when they examine the calculated magnitudes in the correlation matrix, they find that all are low or negligible. "Wait a minute, though!" Just below that low correlation of .2, the computer has printed out that the probability (*p*) is .045. Hey! This correlation is statistically significant at the .05 level! Seeking something of importance to emerge from the investigation, they grasp at this possibility of finding something "significant". The difference between statistical and practical significance must always be kept in mind as explained above with r^2 . The interpretation of a correlation coefficient to describe is different that using it to test a hypothesis.

The null hypothesis tested is $H_0: \rho = 0$. This is saying that the true population parameter (ρ), rho, for the correlation between the variables is zero. Therefore, if a correlation is found to be significant, then this is confirming that the chances that this is an unlucky sample from a population whose true correlation is zero is slight. The real correlation might be (or might not be) .2 in the population. Please note that this in no way indicates that Y is a good or a significant, whatever that might mean, predictor of Y.

However, many researchers who have found "significant" correlations continue by drawing conclusions about and discussing this "significance" as though it had practical importance. Further, no researcher whose research was reviewed in the volumes cited earlier actually posited such hypotheses a priori. Are not agricultural educators just fooling themselves in reporting such significance? This writer would argue that this is just the case. Further, the reporting of these levels of significance, when few understand the hypothesis which has been tested, may tend to further confuse and frustrate the consumers of our research. FFA advisers in local high schools who read that SAE Quality is a "significant predictor" of FFA achievement, when $r = .2$, may be quickly disillusioned with the utility of agricultural education research when they can find no practical relationship in their chapter.

The scope of this problem may in some ways have been aided by the thoroughness of computer print outs. These print outs usually not only report " r ", but " n " and " p ". Had hand calculations been made, few researchers would have looked up the significance level on a table and would have proceeded to use the calculated value for descriptive purposes only. The fact that " p " level is printed on the output makes it all too tempting to report something "significant" when one is seeking just to describe. Further, researchers may not have the computer plot the relationships to check the assumptions of linearity and homoscedasticity as indicated previously. The tendency of agricultural education researchers to report these significance levels may have been partially encouraged by the perception that research out to find something significant or make some new discovery. For an interesting discussion of whether researchers are about making discoveries or eliminating ignorance, see "Ignorance in Educational Research" (Wagner, 1993). Further, the supposition has been forwarded by some that journals will only accept articles in which something "significant" has been found. Herein may lie the philosophical problem of statistical significance versus practical importance. Hopefully, referees for the Journal of Agricultural Education are not so naive as to support a viewpoint that statistical significance must exist in an article for it to be accepted for publication. Historically, the Journal has not seemed to take such a narrow view of science and/or publication policy. Finally, the hypothesis testing example

could be being used just by the naive, neophyte and uninformed.

Summary

Statistics are tools which are used to clarify and aid understanding and not to confound and confuse. The use of hypothesis testing with correlations when the purpose of the research is description seems to only serve the latter purpose. Questions, objectives and hypothesis posed for research in agricultural education most often seek to describe. How often have you seen a study in agricultural education where one of the major objectives was to determine if the correlation between two variables in the population was zero? If inference is to be made to a population, then hypothesis testing is essential. Correlations are an important tool to help us understand whether or not two (or more) variables vary, together, i. e., they help us describe and explain. The profession needs to be able to more fully explain many phenomena operant within our discipline. Description is our most common usage for correlations, but if used for inference, then researchers should demonstrate a full understanding of their usage. Additionally, researchers should not confuse statistical significance with practical significance. Let us not bring confusion by clarity to the readers of our research.

References

- Davis, J.A. (1971). Elementary survey analysis. Englewood, NJ: Prentice-Hall.
- Ferguson, G.A. (1976). Statistical analysis in psychology & education. New York: McGraw-Hill Book Company.
- Wagner, J. (1993, June-July). Ignorance in educational research: Or, How can you not know that? Educational Researcher, 22(5), 15-23.