Mid-Western Educational Researcher

Volume 9
Issue 4 *Regression Analyses*

Article 9

1996

Categorical or Continuous Interaction?

Keith McNeil
New Mexico State University

Isadore Newman
The University of Akron

Categorical or Continuous Interaction?

Keith McNeil, New Mexico State University
Isadore Newman, The University of Akron

Abstract

McClelland and Judd (1993) concluded that many interactions have been found with categorical variables, but few with continuous variables. Using a mathematical point of view, they concluded that investigating interactions with continuous variables was less powerful than with categorical variables. This article analyses the issue from three other points of view: design of the study, measurement of the independent variable, and nature of the question asked. Our present conclusion is that the choice of either categorical or continuous interaction depends upon the research hypothesis being posed and the desired conclusions.

While many statistics texts and researchers view interaction as either a bothersome assumption or a mettlesome outcome clouding the interpretation of main effects, some researchers realize that interaction may be an interesting phenomenon in its own right. Those who do not believe in panaceas argue that what works best may depend upon certain other factors -- implying the need for interaction.

McClelland and Judd (1993) discussed two major kinds of interactions--those resulting from categorical data and those resulting from continuous data. They pointed out that few researchers have obtained interactions with continuous variables, and they presented a mathematical rationale for why this is the case.

While we do not disagree with their mathematics, the purpose of this paper is to point out the differences between categorical interaction and continuous interaction. We do this by focusing on three areas: (a) design differences, (b) measurement differences, and (c) nature of the interaction research hypothesis.

Design Differences

The basic design difference is how the independent variable is conceptualized. The independent variable can be conceptualized as either distinct levels (categorical) or as a continuum. *Categorical* variables are usually studied in controlled situations wherein there is maximum control over the nature of the levels--often maximizing the differences between the levels (as in studies using treatment and control groups). Results from such studies can be generalized to those levels, but not to other levels, not even to levels between the studied levels. Categorical designs are often set up such that the independent variables are uncorrelated, allowing for the additive partitioning of the sum of squares.

Continuous independent variables are usually obtained in a field setting and therefore lack the rigor of laboratory control. Continuous variables do allow for the generalization of results to values between those actually sampled. Continuous variables are usually correlated, but this is the way the real world is and thus how the real world should be modeled.

Measurement Differences

Categorical variables can result from either an inherently nominal variable or an arbitrary categorization of an inherently continuous variable. A continuum may be artificially dichotomized, trichotomized, etc. The limits of these categories are usually determined from the data (such as a median split).

Arbitrary splits in the data limit the comparing of results from one study to another. For example, the numerical value of the median split in one study is unlikely to be the same as in another study. With categorical data, the generalizations are limited to the specific levels that were investigated, whereas the generalizations with continuous data can be made to values between the actual numbers observed.

Continuous variables should be used when one can assume an underlying continuum. Continuous variables allow for smooth continuous relationships to be identified, whereas categorical variables allow only for differences (not even stair-step kinds, unless *a priori* trend analyses are performed). Once a smooth relationship is found, further study must identify if that relationship represents the true functional relationship, or how the variables are scaled.

Identification of the true functional interaction will likely not be easy. But the interactions are likely to be continuous, and the relationship will be a function of how the construct is measured. The GLM approach facilitates investigation of various types of interactions (McNeil, Newman, & Kelly, 1996), as will be shown in a later section.

Nature of the Interaction Research Hypothesis

Categorical interaction is defined as "the differences are different." As a consequence, the test is usually nondirectional, simply looking for any interaction whatsoever. Categorical interaction research hypotheses are usually tested to make sure that the assumption of no interaction is tenable so that the main effects can be interpreted. That is, interaction is usually not viewed as an interesting phenomenon in its own right.

In addition, all possible interactions are usually tested simultaneously. This can be determined from the degrees of freedom associated with the test (4 degrees of freedom in Table 1). The technique of planned comparisons can be extended to the interaction question, but seldom is. For instance, trend effects could be planned ahead of the data collection, and could be the focus of the study instead of the main effects being the focus. For instance, in a two-treatment, pretest-posttest design, the treatment-by-time interaction is the only interesting hypothesis. Both the time and treatment main effects are smaller than what might otherwise be

Table 1.

Source Table for the Data in Figure 1, with Specific Trend Interactions

Source	SS	df	MS	F	p
G	0.00	4	0.00	0.00	1.00
T	0.00	1	0.00	0.00	1.00
G*T	60.00	4	15.00	1500.00	0.0001
G1* T	60.00	1	60.00	6000.00	0.0001
G2* T	0.00	1	0.00	0.00	1.00
G3* T	0.00	1	0.00	0.00	1.00
G4* T	0.00	1	0.00	0.00	1.00
ERROR	0.20	20	.01		

expected--because of the expectation that the two groups are similar at pretest.

Continuous. The definition of interaction with continuous variables is how one variable effects the criterion variable, depending on the value of another continuous variable--the multiplicative effect of two predictors on the criterion. Because some thought is put into the test, the research hypothesis is usually directional, that is one expects high scores on one variable to have a catalytic effect on the criterion with high scores on the other variable. Thus, this catalytic effect is expected, and the researcher hopes to discover it.

Figure 1.

Data analyzed in Study

G	T	D	G	T	D	G	T	D
1	1	2.1	1	1	1.9	1	1	2
2	1	3.1	2	1	2.9	2	1	3
3	1	4.1	3	1	3.9	3	1	4
4	1	5.1	4	1	4.9	4	1	5
5	1	6.1	5	1	5.9	5	1	6
1	2	6.1	1	2	5.9	1	2	6
2	2	5.1	2	2	4.9	2	2	5
3	2	4.1	3	2	3.9	3	2	4
4	2	3.1	4	2	2.9	4	2	3
5	2	2.1	5	2	1.9	5	2	2

Numerical Example

Figure 1 contains fictitious data for a 2 x 5 design. The top part of Table 1 is the source table resulting from these data. Note that the G*T interaction source having 4 degrees of freedom is significant. The bottom part of Table 1 illustrates the source table when all possible interaction trends are analyzed. Note that all of the sum of squares is attributed to the interaction between the "linear trend" in G and T. The global interaction term in Table 1 lumped together all the interaction sums of squares and produced a global test of the interaction. Since the F is significant, the conclusion is that there is interaction somewhere, but the "where" cannot be determined. What is also possible, of course, is that the global test may lead to no significance, while a specific source could be significant.

Researchers who analyze continuous data usually only test the linear interaction--the G^1*T interaction in Table 1. The data in Figure 1 can be treated as continuous by assigning numbers to the levels. We have assigned a "1" to the data in level 1 of G, a "2" to level 2 of G, a "3" to level 3 of G, a "4" to level 4 of G, and a "5" to level 5 of G (and a "1" to level 1 of T and a "2" to level 2 of T). Thus we have two continuous variables, G and T, as well as the trend interactions. Figure 2 contains the General Linear Model approach for testing the linear interaction. The G^1*T linear interaction is tested by comparing the Full Model in part A with the Restricted Model in part A. Notice that the F and probability results are the same as in Table 1. Part B of Figure 2 presents a slightly different approach, testing the linear interaction over and above the linear effects of the two variables. Part C indicates another way to test the interaction--which will be discussed later. Because these three approaches use different models, they are actually testing slightly different research questions. One would not want to make the error of incorrectly testing the "interaction question"--a type VI error as discussed by Newman, Deitchman, Burkholder, Sanders, and Ervin (1976), and Newman and Newman (1994).

Figure 2.
Various GLM Analyses of the interaction hypothesis.

A. Replication of the categorical approach

Full Model: $D = a*U + t*T + g1*G^1 + g2*G^2 + g3*G^3 + g4*G^4 + i1*G^1*T + i2*G^2*T + i3*G^3*T + i4*G^4*T + E1$
Want $i1 < 0$ Restrict $i1 = 0$

Restricted Model: $D = a*U + t*T + g1*G^1 + g2*G^2 + g3*G^3 + g4*G^4 + i2*G^2*T + i3*G^3*T + i4*G^4*T + E2$

Numerator MS:60.00	df: 1	F value: 6000.00
Denominator MS:0.01	df: 20	prob: 0.0001

B. Testing linear interaction over and above the two linear effects.

Full Model: $D = a*U + t*T + g*G + i*G^1*T + E3$
Want $i < 0$ Restrict $i = 0$

Restricted Model: $D = a*U + t*T + g*G + E4$

Numerator MS: 60.00	df: 1	F value: 7800.00
Denominator MS: 0.007692	df: 26	prob: 0.0001

C. Testing the interaction component.

Full Model: $D = a*U + i*G*T + E5$
Want $i < 0$ Restrict $i = 0$

Restricted Model: $D = a*U + E6$

Numerator MS: 60.00	df: 1	F value: 8400.00
Denominator MS: 0.0071428	df: 28	prob: 0.0001

Relying on the trend notions discussed above, other kinds of interactions can be studied. One would want to investigate such interactions for one of two reasons. Either the true functional relationship between the independent variables and dependent variable is that way, or there is a scaling problem with one or both independent variables, and the nonlinear interaction maps the scaling problem.

Just as with the categorical variables in Part C of Figure 2, one interaction term between continuous variables may explain all the variance. McNeil (1970) illustrated how Newton might have obtained the Law of Gravity, which contains only one interaction term, and no main effects terms. The interaction term is between the linear component of gravity and the squared value of time ($D = 1/2\ G*T^2$).

Conclusion

Clearly different interaction questions are being asked with categorical and continuous data. Since all of the interactions are lumped together in the categorical approach, one would expect to more often find significant interactions with the categorical approach. But what the researcher does with those interactions (usually ignores them and looks at simple effects) and the extent to which the researcher is really interested in finding them (blast it, muddies up my main effects again!) is probably more important than whether or not they are found. Replication of unexpected interactions is another issue that should be considered, but unfortunately is not. If you don't expect them and you don't want them, why would you even consider replicating the research to find out if they are a stable phenomenon?

Testing for "interaction" is, after all, testing a specific hypothesis, and that should never be forgotten--but it often is. If two statistical tests are testing different questions, then they should not be compared for their relative power. What they should be compared for is the reasonableness of the research question that is being tested. If one is only interested in making sure that the assumption of no interaction is a reasonable assumption, then the use of categorical interactions makes sense. If one is interested in determining functional relationships, then the use of continuous interactions makes sense.

References

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, 114(2), 376-390.

McNeil, K. (1970). Meeting the goals of research with the multiple regression analysis. *Multivariate Behavioral Research*, 5, 375-386.

McNeil, K., Newman, I., & Kelly, F. J. (1996). *Testing research hypotheses with the general linear model.* Carbondale, IL: Southern Illinois University Press.

Newman, I., Deitchman, R., Burkholder, J., Sanders, R., & Ervin, L. (1976). Type VI error: Inconsistencies between the statistical procedure and the research question. *Multiple Linear Regression Viewpoints*, 6(4), 1-19.

Newman, I., & Newman, C. (1994). *Conceptual statistics for beginners* (2nd edition). Lanham, MD:University Press of America.