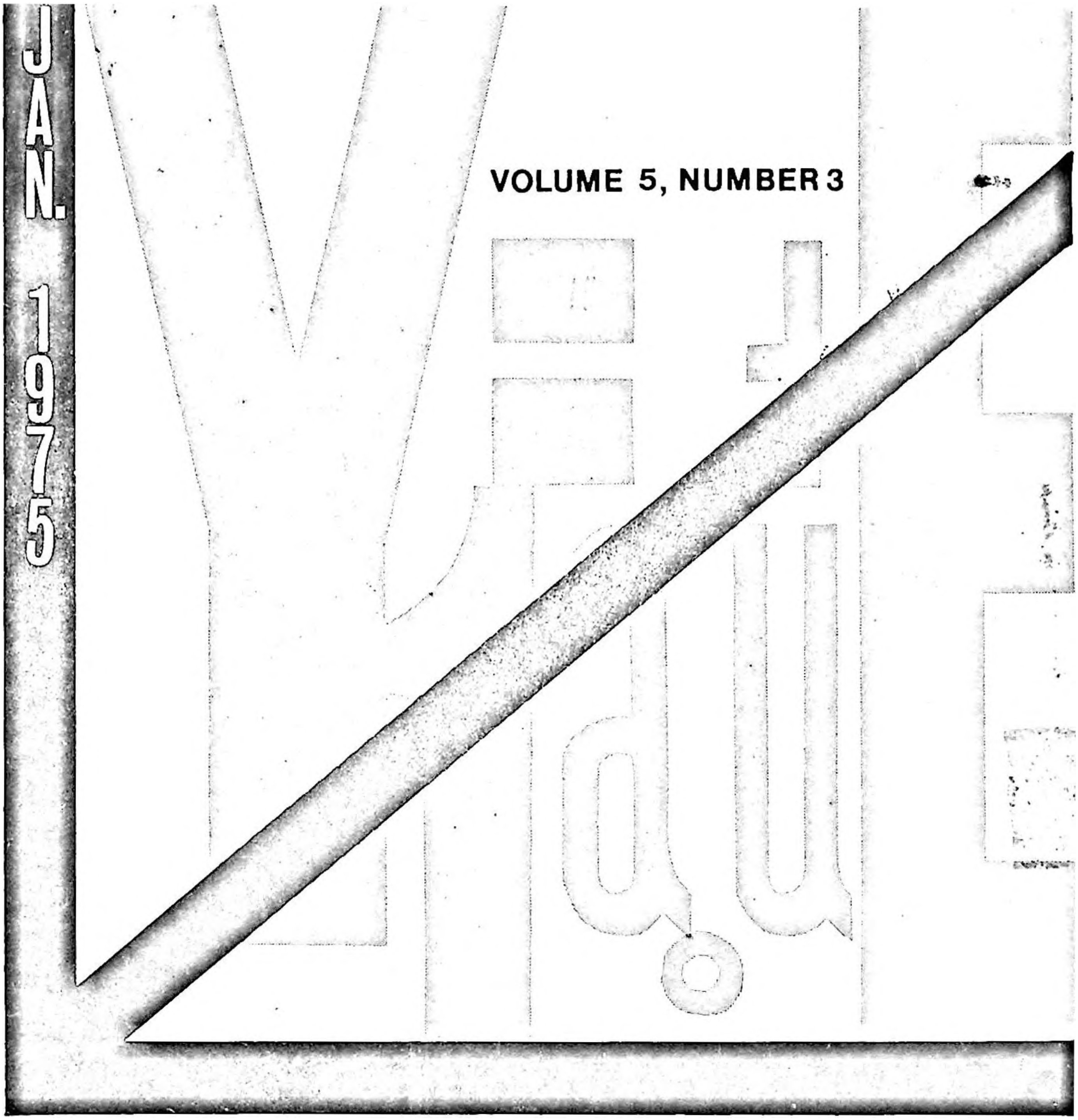

MULTIPLE LINEAR REGRESSION VIEWPOINTS

A publication of the Special Interest Group
on Multiple Linear Regression
of The American Educational Research Association.

IF YOU ARE SUBMITTING A RESEARCH ATRICLE OTHER THAN NOTES OR COMMENTS, I WOULD LIKE TO SUGGEST THAT YOU USE THE FOLLOWING FORMAT, AS MUCH AS POSSIBLE:

TITLE
AUTHOR AND AFFILIATION
SINGLE-SPACED INDENTED ABSTRACT (THE REST OF THE MANUSCRIPT SHOULD BE DOUBLE-SPACED)
INTRODUCTION (PURPOSE—SHORT REVIEW OF LITERATURE, ETC.)
METHOD
RESULTS
DISCUSSION (CONCLUSION)
REFERENCES

ALL MANUSCRIPTS SHOULD BE SENT TO THE EDITOR AT THE ABOVE ADDRESS. (ALL MANUSCRIPTS SHOULD BE CAMERA-READY COPY)

TABLE OF CONTENTS

Multiple Linear Regression Viewpoints
Vol. 5, No. 3, 197-

PATH ANALYSIS AND CAUSAL MODELS AS REGRESSION TECHNIQUES

John D. Williams
University of North Dakota

and

Ronald J. Klimpel
North Dakota Combined Law En orcement Council

Summary

Path analysis and causal models, two closely related concepts, are explained in detail from a regression viewpoint. An example from land tenure and community leadership is given in relationship to the recursive equations; the predictions of partial coefficients to vanish (i.e., equal zero) is seen to be analogous to the dropping of the corresponding partial regression coefficient from the linear model. The meaning of direct effect is explored and contrasted to a more common regression concept, the independent contribution of a variable. The reproduction of the correlation matrix under the restriction that one or more paths have been dropped is also shown.

The concept of path analysis originated in the work of Wright (1921, 1934, 1954) in relation to developing genetic models. A closely related concept uses the idea of causal models, with the major contributions to this area being Simon (1954, 1957) and Blalock (1962, 1964, 1967, 1971), with Blalock's contribution being in sociology. A closely related approach in economics was given by Wold and Juneen (1953) and Wold (1954). Two descriptions of path analysis within the area of sociology were made by Land (1969) and Heise (1969). More recently, Kerlinger and Pedhazur (1973) described the path analysis-causal models concept in a regression framework.

Because the path analysis-causal models concept is easily translated into a rather straightforward set of regression equations, it seems reasonable to describe path analysis in regression terms and to point out some of the areas of dispute within the various contributions to this area. To do so, a rather simple application is given, and the model is then described from the several points of view.

An Example From Land Tenure and Community Leadership

Klimpel (1974) was interested in finding the relationship that existed between land ownership and the community leadership of people who lived in a unique sociological condition in an isolated North Dakota area. Because of the construction of a dam on the Missouri River, people who lived in three small communities in close proximity were required to move, and most of the affected people moved to a central community in 1955. Klimpel was interested in finding what relationships existed prior to 1955 in regard to land ownership and community leadership, and what influence these two variables had on present community leadership. Additionally, he wanted to see what relationships existed between prior (pre 1955) land ownership and present land ownership, and also what relationship existed between present land ownership and present community leadership.

Klimpel defined land ownership in terms of number of acres owned; an interview schedule was developed to discern community leadership. A sample of adults were asked to list up to five community leaders in the following five areas for both past and present: general leader, leader in special projects, leader at the state level, leader at the national level, and covert community leader. A total was then found separately for the number of mentionings both past and present.

There were then four variables of interest:

X_1 = ownership in acreage in the past;

X_2 = number of mentionings for leadership in the past;

X_3 = ownership in acreage (present);

X_4 = number of mentionings for leadership in the present.

Typically, a set of recursive equations are used to define the relationship between the four variables:

$$X_2 = a_1 + b_{21}X_1 + e_1, \qquad (1)$$

$$X_3 = a_2 + b_{31.2}X_1 + b_{32.1}X_2 + e_2, \text{ and} \qquad (2)$$

$$X_4 = a_3 + b_{41.23}X_1 + b_{42.13}X_2 + b_{43.12}X_3 + e_3. \quad (3)$$

In the recursive equations, each variable X_i is considered to be a possible cause of the variables occurring after X_i but not before X_i. The values a_1 - a_3 represent the intercepts for the corresponding equations. The regression coefficients given in equations 1-3 are indicated as partial regression weights. If a path diagram were made of the recursive equations, Figure 1 would result.

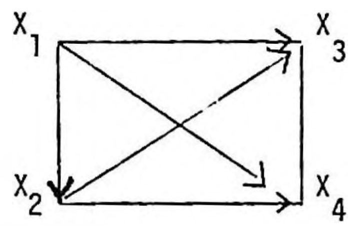

Figure 1. Recursive Path Diagram.

Quite often, the interest in path analysis is not in a complete set of paths as given by Figure 1, but rather in restricting the number of paths. For the land tenure data, the path between X_2 and X_3 was felt to be unneeded; that is, it did not seem likely that ownership in land in the present was caused by community leadership in the past, controlling for previous land ownership. Dropping the path for X_2 and X_3 would then transform equation 2 to be

$$X_3 = a_4 + b_{31}X_1 + e_4. \quad (4)$$

This equation is equivalent to saying $b_{32.1}$ vanishes (or is equal to zero). There are several ways to test the hypotheses suggested by the absence of a path from X_2 to X_3. Blalock (1964) shows that the vanishing of $b_{32.1}$ is equivalent to the hypothesis $r_{32.1} = 0$. However, if equation 2 is seen as a full model and equation 4 is seen as a restriction of equation 2, then the significance of a path between X_2 and X_3 can be tested by

$$F = \frac{(R^2_{3.12} - R^2_{31})/1}{(1 - R^2_{3.12})/(N - k - 1)}, \qquad (5)$$

where k is the number of paths to X_3 when all paths are represented.

Also, most users of path analysis (notably excepting Blalock) focus on the standardized regression weights. The standardized regression weights (beta coefficients) are called path coefficients. Using Klimpel's data (N = 57), equations 1, 3 and 4 are represented with path coefficients and correlation coefficients in Figure 2.

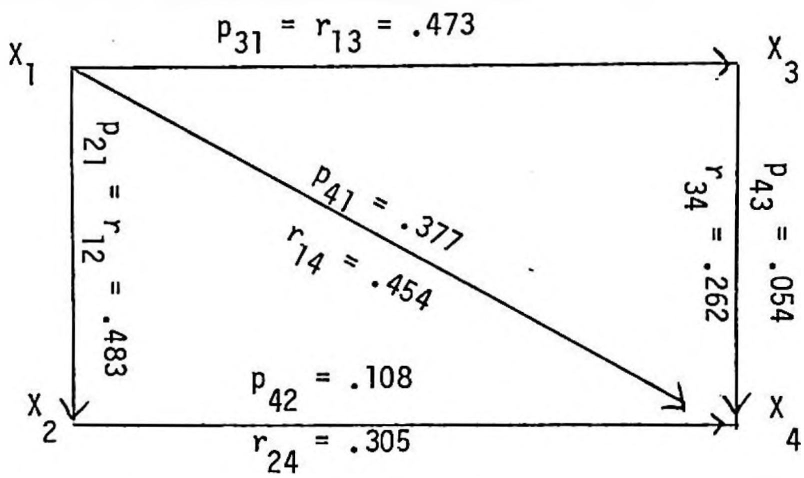

Figure 2. Path Diagram for Land Tenure Data with $p_{23} = 0$.

Several important points can be made regarding Figure 2. One correlation coefficient has not been included; $r_{23} = .275$. The hypothesis $r_{23.1} = 0$ is tenable, as $r_{23.1}$ can be found to be .060. Also, the test given by equation 5 yields

$$F = \frac{(.227 - .224)/1}{(1 - .227)/54} = .210,$$ showing that the path from X_2 to X_3 can be dropped without losing a significant amount of information regarding X_3.

Direct and Total Indirect Effects in Path Analysis

An important principle of path analysis is that the path coefficients measure the direct effect of one variable upon the other; the difference between the correlation coefficient and the path coefficient is the total indirect effect.

Total Indirect Effect of Variable i of Variable j = $r_{ij} - p_{ji}$. (6) Also, the juxtaposition of i and j with the path coefficient is deliberate. With path coefficients, the variable given by the first number (j) is "caused" by the variable given by the second number (i).

For the four variables given in Figure 2, the total indirect effect (TIE) of each variable can be found and are shown in Table 1.

TABLE 1
The Total Indirect Effects for the Data in Figure 2

TIE of 1 on 2 = $r_{12} - p_{21}$ = .483 - .483 = 0

TIE of 1 on 3 = $r_{13} - p_{31}$ = .473 - .473 = 0

TIE of 1 on 4 = $r_{14} - p_{41}$ = .454 - .377 = .077

TIE of 2 on 4 = $r_{24} - p_{42}$ = .305 - .108 = .197

TIE of 3 on 4 = $r_{34} - p_{43}$ = .262 - .054 = .208

Several comments are necessary regarding the results in Table 1. In that no arrows except variable 1 are shown to be causal to variables 2 and 3, there can be no indirect effect of variable 1 on either of variables 2 and 3. The TIE of variable 1 on variable 4 is fairly minimal. This implies that ownership in acreage in the past (X_1) makes a substantial direct effect on the number of mentionings for leadership in the present (X_4). On the other hand, the total indirect effect of variables 2 and 3 on 4 are fairly substantial; the size of the path coefficients together with the indirect effects of each variable would tend to indicate that the strength of p_{42} and p_{43} is such that they might be dropped from the analysis. The implication of this conjecture is that the number of mentionings for leadership in the past (X_2) and ownership in acreage in the present (X_3) have only a weak direct relationship with the number of mentionings for leadership in the present (X_4).

Before a reanalysis is performed dropping p_{43} and p_{42}, a comment is necessary regarding the concept of direct and indirect effect; the terms should not be confused with the more familiar measure, the independent contribution of a variable. The squared path coefficient is said to measure the proportion of the variance of the dependent variable for which the determining variable is directly responsible (Wright, 1934). While it would seem the independent contribution of a variable means the same thing as the proportion of the variance of the dependent variable for which the determining variable is directly responsible, it can be shown empirically that these two concepts are not identical. Table 2 contains both the independent contribution of each variable in explaining the variance in variable 4 and the squared path (beta) coefficients.

TABLE 2
Comparison of Squared Path Coefficients and Independent Contribution for Land Tenure Data

	INDEPENDENT CONTRIBUTION	SQUARED PATH COEFFICIENT
Variable 1:	$R^2_{4,123} - R^2_{4,23} =$ $.218 - .127 = .091$	$P^2_{41} = .377^2 = .142$
Variable 2:	$R^2_{4,123} - R^2_{4,13} =$ $.218 - .209 = .009$	$P^2_{41} = .108^2 = .012$
Variable 3:	$R^2_{4,123} - R^2_{4,12} =$ $.218 - .216 = .002$	$P^2_{43} = .054^2 = .003$

Table 2 is important for showing what a path coefficient (or path coefficient squared) is <u>not</u>. The difficulty is in assessing what is meant by the term "directly" in the statement, "the squared path coefficient measures the proportion of the variance of the dependent variable for which the determining variable is <u>directly</u> responsible." While the left half of Table 2 is presently called independent contribution, it has been variously described by the following names: unique contribution or usefulness (Darlington, 1968). However, the term <u>directly</u> responsible is not synonymous with the just mentioned terms. Rather, what is meant by direct, the term "directly responsible" is given by the following equation:

$$R^2 = p_{y1}^2 + p_{y2}^2 +\ldots+ p_{yk}^2 + 2p_{y1}p_{y2}r_{12} + 2p_{y1}p_{y3}r_{13} +\ldots+ 2p_{y(k-1)}p_{yk}r_{(k-1)k}, \quad (7)$$

where k is the number of predictor variables preceding y in the path analysis that have an arrow directly to y. Equation 7 is due to Englehart (1936). If equation 7 is applied to the four variable models for the land tenure data,

$$R^2_{4.123} = p_{41}^2 + p_{42}^2 + p_{43}^2 + 2p_{41}p_{42}r_{12} + 2p_{41}p_{43}r_{13} + 2p_{42}p_{43}r_{23}, \quad (8)$$

and

$$R^2_{4.123} = .142 + .012 + .003 + .039 + .019 + .003 = .218.$$

Thus, the squared path coefficients are said to make a <u>direct</u> contribution to the criterion variance; an additional $\binom{k}{2} = \frac{(k)(k-1)}{2}$ terms are said to make a <u>joint</u> contribution; for equation 8, the sum of the three squared path coefficients is .157, and the joint contribution of the three variables is .061, totaling $R^2_{4.123}$.

An Alternative Approach to Path Analysis

If the variables are ordered in the same manner as was true for the recursive equations, then an alternative manner to report the results of the path analysis seems worthy of consideration. First, three equations are necessary:

$$X_4 = a_5 + b_{41} X_1 + e_4, \qquad (9)$$

$$X_4 = a_6 + b_{41.2} X_1 + b_{42.1} X_2 + e_5, \qquad (10) \quad \text{and}$$

$$X_4 = a_3 + b_{41.23} X_1 + b_{42.13} X_2 + b_{43.12} X_3 + e_3. \qquad (11)$$

Equation 11 is identical to equation 3.

Now, for correlated data,

$$R^2_{y.123...k} = r^2_{y1} + r^2_{y(2.1)} + ... + r^2_{y(k.k-1...1)}, \qquad (12)$$

which, for four variables is given by

$$R^2_{4.123} = r^2_{41} + r^2_{4(2.1)} + r^2_{4(3.12)}. \qquad (13)$$

Equations 11 and 12 require the same ordering as in the recursive equations, but allow an assessment of the additional predictability at each new path. Thus, since X_1 is the first variable, r^2_{41} is found. Then, X_2 is adjusted for X_1, and a part correlation is found between the unmodified criterion (X_4) and X_2, adjusted for X_1. Finally, variable 3 is adjusted for both variables 1 and 2 and then correlated with X_4.

Actually, it is not necessary to calculate the part correlations directly, or employ additional specialized computer programs; the values for $r^2_{4(2.1)}$ and $r^2_{4(3.12)}$ can be found directly from the use of equations 9, 10 and 11. From equation 9, $R^2_{41} = .206$, from equation 10,

$R^2_{4.12} = .216$, and from equation 11, $R^2_{4.123} = .218$. Thus, $r^2_{4(2.1)} = .216 - .206 = .010$, and $r^2_{4(3.12)} = .218 - .216 = .002$. Also, $r_{4(2.1)} = .10$ and $r_{4(3.2)} = .045$.

The usefulness of the formulation given by equations 12 and 13 is that the path analysis can be thought of in terms of a logical ordering: a first variable is determined, and each subsequent variable included in the analysis is adjusted for earlier variables in the analysis.

A Backward Model Using Part Notation

As can be readily discerned, the formulation given by equations 9-13 can be conceived as a "forward" model of path analysis; the earliest causes are accounted for first, and sequentially the causes are included as they occur in the causal ordering. The process can be conceptually reversed, first assessing most immediate causes, and sequentially including more remote causes using the part correlation approach. For the four model approach, the linear models would be

$$X_4 = a_7 + b_{43}X_3 + e_6, \qquad (14)$$

$$X_4 = a_8 + b_{42.3}X_2 + b_{43.2}X_3 + e_7, \qquad (15) \quad \text{and}$$

$$X_4 = a_3 + b_{41.23}X_1 + b_{42.13}X_2 + b_{43.12}X_3 + e_3. \qquad (16)$$

where equation 16 is identical to equations 3 and 11. Equations similar to those given for equations 12-13 are easily developed:

$$R^2_{y.123...k} = r^2_{yk} + r^2_{y(K-1.K)} + \ldots + r^2_{y(1.23...k)}, \qquad (17)$$

which for four variables is given by

$$R^2_{y.123...k} = r^2_{43} + r^2_{4(2.3)} + r^2_{4(1.23)}. \qquad (18)$$

The calculations for the present data yield

$$R^2_{4.123} = .262^2 + .242^2 + .302^2 = .218,$$

where

$$r_{4(2.3)} = .242 \text{ and } r_{4(1.23)} = .302.$$

Both the forward and backward analyses are suggested by Duncan (1970) as a means of partitioning the variance in path analysis; Duncan does not actually use a part correlation process directly, however.

Towards a More Parsimonious Theory--
Reducing the Number of Paths

While the model described in Figure 2 explicates Klimpel's original hypotheses, no injustice accrues to considering the possibility of dropping negligible paths.

Surely, the smaller the number of paths, the more parsimonious any causal explanation can be. Referring back to Figure 2, it is possible to test each path coefficient for significance. The test for path (beta) coefficients being significantly different from zero is given by the usual t test for partial regression coefficients. Table 3 contains the t tests for each of the five path coefficients.

TABLE 3
Tests of Significance for the Path Coefficients

Path Coefficient	t Value
p_{21}	4.092*
p_{31}	3.983*
p_{41}	2.482*
p_{42}	.777
p_{43}	.390

Since p_{42} and p_{43} are non-significant, a new path analysis containing only the significant paths (p_{21}, p_{31} and p_{41}) can be undertaken. The results are shown for this new analysis in Figure 3.

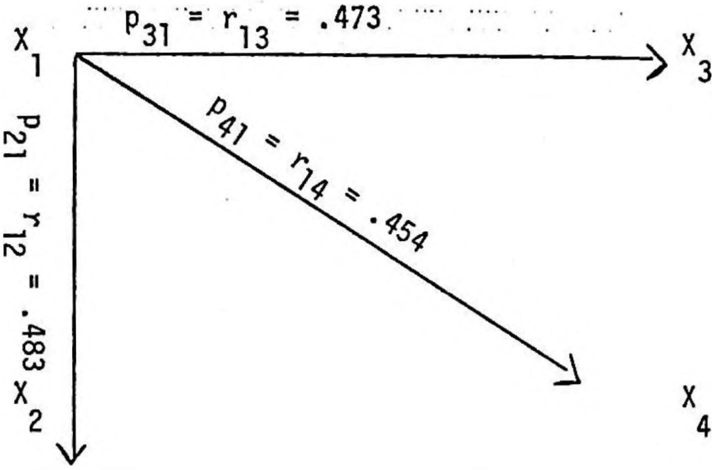

Figure 3. Path Diagram with Three Paths

In that the paths from X_2 to X_4 and X_3 to X_4 have been deleted, it should also be noted that two hypotheses have been tested by the previous non-significant t tests: They are $r_{24.13} = 0$ and $r_{34.12} = 0$. Using Blalock's (1962) approach, the hypotheses to be tested by the model in Figure 3 are $r_{24.1} = 0$ and $r_{34.1} = 0$. If the paths p_{42} and p_{43} are simultaneously dropped from equation 3, the reduced path equation, already given by equation 9, $X_4 = a_5 + b_{41}X_1 + e_7$, contains only p_{41}. When simultaneously dropping p_{42} and p_{43}, only X_1 need be partialled out; the predicted relationships are $r_{24.1} = 0$ and $r_{34.1} = 0$, the same predictions as are given in Blalock. Empirically, $r_{24.1} = .109$ and $r_{34.1} = .060$; both values are reasonably close to zero.

The t test given in Table 3 is not sufficient for testing these two partial regression coefficients separately. First, two additional regression models could be developed to test the significance of $r_{24.1}$ and $r_{34.1}$:

$$X_4 = a_9 + b_{41.2}X_1 + b_{42.1}X_2 + e_8, \qquad (19)$$

and

$$X_4 = a_{10} + b_{41.3}X_1 + b_{43.1}X_3 + e_9. \qquad (20)$$

Using equations similar to equation 5,

$$F = \frac{(R^2_{4.12} - R^2_{41})/1}{(1 - R^2_{4.12})/54} = \frac{(.216 - .206)/1}{(1 - .216)/54} = .689$$

and

$$F = \frac{(R^2_{4.13} - R^2_{41})/1}{(1 - R^2_{4.13})/54} = \frac{(.209 - .206)/1}{(1 - .209)/54} = .205.$$

These F values are the tests of significance for $r_{24.1}$ and $r_{34.1}$ respectively and both are non-significant.

Because both p_{42} and p_{43} are being simultaneously dropped, an additional test is necessary to find whether they can both be dropped without a significant loss of information. The test is similar to equation 5 and is given by

$$F = \frac{(R^2_{4.123} - R^2_{41})/2}{(1 - R^2_{4.123})/53} = \frac{(.218 - .206)/2}{(1 - .218)/53} = .407.$$

Since this F value is non-significant, it can be concluded that p_{42} and p_{43} can be dropped without losing a significant amount of information. Thus, a more parsimonious path analysis is given in Figure 3, using only three paths, p_{21}, p_{31}, and p_{41}. In terms of the land tenure data, a rather simple explanation ensues. For those individuals in the population, land ownership in the present (X_3), number of times mentioned for leadership in the past (X_2) and number of times mentioned for leadership in present (X_4) are caused by past land ownership (X_1); further any relationships that exist between X_2, X_3 and X_4 can be explained in terms of their all being caused by X_1 (past land ownership).

Reproducing the Correlation Matrix

One of the more important contributions of path analysis is the reproduction of the correlation matrix. The correlation matrix can be reproduced using the following equation

$$r_{ij} = \sum_k p_{jk} r_{ik}, \qquad (21)$$

where k includes every variable which has a path leading directly to X_j. For situations such as are depicted in Figure 1 wherein every variable is connected to every other variable with a path, equation 21 would identically reproduce the correlation matrix. On the other hand, models such as Figure 2 and Figure 3 would tend to have some departure from an exact reproduction of the correlation matrix. Figure 2 has p_{32} deleted; thus if equation 21 is used successively,

$$r_{12} = p_{21} r_{11} = p_{21} = .483, \qquad (22)$$

$$r_{13} = p_{31} r_{11} = p_{31} = .473, \qquad (23)$$

$$\begin{aligned} r_{14} &= p_{41} r_{11} + p_{42} r_{12} + p_{43} r_{13} \\ r_{14} &= p_{41} + p_{42} r_{12} + p_{43} r_{13} \end{aligned} \qquad (24)$$

$$r_{14} = .377 + (.108)(.483) + (.054)(.473) = .454,$$ as

was true previously;

$$r_{23} = p_{31} r_{21} = p_{31} r_{12} \qquad (25)$$

$r_{23} = (.473)(.483) = .228.$

Actually, $r_{23} = .275$;

$r_{24} = p_{41}r_{21} + p_{42}r_{22} + p_{43}r_{23},$

$r_{24} = p_{41}r_{12} + p_{42} + p_{43}r_{23},$ (26)

$r_{24} = .377(.483) + (.108) + (.054)(.275) = .305$, as

was true previously; and

$r_{34} = p_{41}r_{31} + p_{42}r_{32} + p_{43}r_{33},$

$r_{34} = p_{41}r_{13} + p_{42}r_{23} + p_{43},$ (27)

$r_{34} = (.377)(.473) + (.108)(.275) + .054 = .262$ as

was true previously. Actually, the only inaccuracy occurred with r_{23}, which was actually .275, but found to be .228 with equation 21. Recalling that p_{32} was not included in this analysis, this slight difference in reproducing the correlation matrix seems tolerable.

If on the other hand the interest is in comparing the reproduced correlation matrix to the path diagram with only three paths (Figure 3), a more difficult test of the system is made.

The equations for r_{12} and r_{13} would be unchanged from equations 22 and 23 in the three path model; hence $r_{12} = .483$ and $r_{13} = .473$.

On the other hand,

$r_{14} = p_{41}r_{11} = p_{41} = .454;$ (28)

$r_{23} = p_{31}r_{21} = p_{31}r_{12}$ (29)

$r_{23} = (.473)(.483) = .228.$ Actually, $r_{23} = .275$;

$r_{24} = p_{41}r_{21} = p_{41}r_{12},$ (30)

$r_{24} = (.454)(.483) = .219$; actually

$r_{24} = .305$.

And finally,

$$r_{34} = p_{41} r_{31} = p_{41} r_{13} \qquad (31)$$

$r_{34} = (.454)(.473) = .215$; actually, $r_{34} = .262$.

Table 4 contains the actual correlations in the portion above the main diagonal in the correlation matrix and the projected correlations under the reduced paths are found below the main diagonal.

TABLE 4
Actual (Above Main Diagonal) and Projected (Below Main Diagonal) Correlations for Land Tenure Data

	X_1	X_2	X_3	X_4
X_1		.483	.473	.454
X_2	.483		.275	.305
X_3	.473	.228		.262
X_4	.454	.219	.215	

Legend:

X_1 = ownership in acreage in the past;

X_2 = number of mentionings for leadership in the past;

X_3 = ownership in acreage (present);

X_4 = number of mentionings for leadership in the present.

Table 4 thus is illustrative of the reasonably close relationship between the actual and projected correlations under the restriction that p_{32}, p_{42}, p_{43} are all zero, indicative that little information is lost by the non-inclusion of these three paths.

Exogenous and Endogenous Variables

To this point, an effort has been mede to present path analysis from a regression framework, using much of the terminology more familiar to the regression formulations. Thus, many of the niceties of path analysis have been omitted. To compensate for this regression emphasis, many of the concepts more familiar to path analysis view are now considered.

Variables in a path analysis are either exogenous or endogenous. While many authors have described these two terms in considerable detail, one major difference is this: a variable that serves only as an independent (or predictor) variable is considered to be exogenous; exogenous variables are often thought to be pre-existing to a system implied in a path analysis. Accordingly, while an endogenous variable may be an independent variable in part of the system, it must be a dependent variable at least once in the system. For the land tenure problem, prior land ownership is considered to be an exogenous variable, with the remaining three variables being endogenous. On the other hand, had some other variable, such as age in 1950 (X_0) been included as a predictor of X_1, then X_0 would also be endogenous. While for many of the sociological examples the explication of exogenous variables may be quite meaningful, this is not always the case.

The usual practive is to consider that all endogenous variables are measured with error, but exogenous variables are measured without error. The usual system in path analysis then includes a residual predictor variable for each endogenous variable, together with an appropriate path coefficient, that accounts for all the remaining variance not attributable to the system. That custom was not followed in this paper; the interested reader is directed to Land (1969) regarding this aspect of path analysis. However, path coefficients for residual variables are easily found as $\sqrt{1 - R^2}$ where R is the multiple correlation of the particular variable with all of it's predictor variables. The residual path coefficient is more commonly referred to as the coefficient of alienation in elementary statistics texts.

Path Coefficients vs. Path Regression Coefficients

This paper has emphasized the use of path coefficients (standardized beta weights). A lively debate has occurred in path analysis regarding using path coefficients or path regression coefficients (non-standardized partial regression weights). Most of the earlier applications of path analysis followed Wright's usage of path coefficients. Principally due to Tukey's (1954) argument for the path regression coefficients, a parallel development of path regression coefficients in path analysis has occurred (for example, see Turner and Stevens, 1959). Tukey (1954, p. 45) has said that, "It is probably unwise to try to assign relative determination to correlated determining variables." This would lead him to show little interest in finding such things as a Total Indirect Effect, which would consequently lessen his interest in path coefficients per se. More importantly, the use of path regression coefficients comes closer to a goal of science: to establish coefficients that are relatively stable and can be used directly in a prediction equation. Wright (1960) has re-argued the case for path coefficients, emphasizing the simplicity of their form; Wright sees the two kinds of coefficients as complimentary rather than as competing concepts.

Empirically, most computer programs yield both kinds of coefficients; even if a researcher has opted for one type of coefficient over the other type, a formulation in the alternative type coefficient would be available.

Discussion

The present paper has as it's main focus to present some of the more simple ideas regarding path analysis from a regression viewpoint. Accordingly, the presentation has been made from the point of view of directing the attention to a regression solution rather than retain the arguments and intentions of the proponents of path analysis. Thus, several concepts have been presented from the point of view of effecting a solution rather than retain the thinking of path analysis. As an example of an omission of the path analysis viewpoint in favor of a pragmatic solution, equation 21 was used for reproducing a correlation matrix. Users of path analysis, on the other hand, are also interested in a complex analysis of the interrelations of the path coefficients and correlation coefficients in any given correlation among endogenous variables. The results from any given calculation from equation 21 could be extended by substitution until the only correlations that appear in the final equation are correlations among exogenous variables. For example, from equation 27,

$$r_{34} = p_{41}r_{13} + p_{42}r_{23} + p_{43}.$$

But $r_{13} = p_{31} + p_{32}r_{21}$, (32)

$$r_{12} = p_{21}, \quad (33)$$

and

$$r_{23} = p_{31}r_{12}, \quad (34)$$

where equations 32, 33 and 34 were each found by successive application of equation 21. Then, by substitution,

$$r_{34} = p_{41}(p_{31} + p_{32}r_{12}) + p_{42}(p_{31}p_{21}) + p_{43},$$

$$r_{34} = p_{41}p_{31} + p_{41}p_{32}p_{21} + p_{42}p_{31}p_{21} + p_{43}. \quad (35)$$

Then, $r_{34} - p_{43} = p_{41}p_{31} + p_{41}p_{32}p_{21} + p_{42}p_{31}p_{21}$ (36)

is a measure of the total indirect effect given earlier. Equation 36 is said to help understand better the variables as they make up the indirect effect of a variable. While this may well be true, the amount of algebra involved easily becomes excessive.

Other issues have been underplayed also. The concept of under and over identification of the estimates of the coefficients is one such issue. While no specific mention was made of it, whenever a path is deleted, an implication is that overidentification is present, and will allow for the testing of the adequacy of the path model. Other concerns of path analysis, including reliability, unmeasured variables, categorical variables and multi-trait-multi-method matrices are treated by Werts and Linn (1970) and are not repeated here.

References

Blalock, H. M. Four-variable causal models and partial correlation. American Journal of Psychology, 1962, 68, 182-194.

Blalock, H. M. Causal inferences in nonexperimental research. Chapel Hill: University of North Carolina Press, 1964.

Blalock, H. M. Path coefficients versus regression coefficients. American Journal of Sociology, 1967, 72, 675-676.

Blalock, H. M. (Ed.). Causal models in the social sciences. Chicago: Aldine, 1971.

Darlington, R. B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.

Duncan, O. D. Partials, partitions and paths. In Borgatta, E. F., and Bohrnstedt, (Eds.). Sociological Methodology 1970. San Francisco: Jossey-Bass, 1970.

Englehart, M. D. The technique of path analysis. Psychometrika, 1936, 1, 287-293.

Heise, D. R. Problems in path analysis and causal analysis. In Borgatta, E. F., and Bohrnstedt, (Eds.). Sociological Methodology 1969, San Francisco: Jossey-Bass, 1969.

Kerlinger, F. N., and Pedhazer, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart and Winston, 1973.

Klimpel, R. J. Longitudinal analysis of community power structures in a rural North Dakota community and it's trade area. Unpublished Master's thesis, The University of North Dakota, 1974.

Land, K. C. Principles of path analysis. In Borgatta, E. F., (Ed.), Sociological Methodology 1969. San Francisco: Jossey-Bass, 1969.

Simon, H. A. Spurious correlations: A causal interpretation. Journal of the American Statistical Association, 1954, 49, 467-479.

Tukey, J. W. Causation, regression and path analysis. In Kempthorne, O., Bancroft, T. A., Gowen, J. W., & Lush, J. D., (Eds.). Statistics and Mathematics in Biology. Ames: Iowa State College Press, 1954.

Werts, C. E., and Linn, R. L. Path analysis: Psychological examples. Psychological Bulletin, 1970, 74, 193-212.

Wold, H. Causality and econometrics. Econometrics, 1954, 22, 162-177.

Wold, H., and Juneen, L. Demand analysis. New York: Wiley, 1953.

Wright, S. Correlation and causation. Journal of Agricultural Research, 1921, 20, 557-585.

Wright, S. The method of path coefficients. Annals of Mathematical Statistics, 1934, 5, 161-215.

Wright, S. The interpretation of multivariate systems. In Kempthorne, O., Bancroft, T. A., Gowan, J. W., & Lush, J. D., (Eds.). Statistics and Mathematics in Biology. Ames: Iowa State College Press, 1954.

Wright, S. Path coefficients and path regressions: Alternative or complimentary concepts. Biometrics, 1960, 16, 189-202.

Vol. 5, No. 3, 1974

FROM EDUCATIONAL EVALUATION TO DECISION MAKING: JAN TO THE RESCUE

Samuel R. Houston
University of Northern Colorado

and

E. LaMonte Ohlson
The Ohio State University

The life of every individual and organization abounds with decisions made in the face of uncertainty. The application of mathematics and statistics to theories of decision-making has helped to examine some of the possible ways by which man arrives at his rational conclusions. Researchers have used game theory, information theory, linear programming, and statistical theory to formulate rules as substitutes for people making decisions about inventory control, production scheduling, quality control, long range allocation of resources, etc. It may be said to be the purpose of this paper is to describe a decision-capturing procedure, Judgment Analysis, which can make explicit, and therefore subject to logical inquiry, how administrators or decision makers (judges) make decisions under a variety of conditions and to suggest ways in which the technique might be extended to the area of evaluating instructional programs and research proposals.

1. Judgment Analysis--General Applications

Judgment ANalysis (JAN) has already been demonstrated to be an extremely useful methodology for describing and captur- the strategies of raters with respect to the responses to a set of complex stimuli. JAN has been used in several studies conducted by the U. S. Air Force for job evaluations and to simulate officer promotion boards with a high degree of efficiency. Equations have also been designed to simulate career counselors in making initial assignments of airmen graduating from basic training (6).

The JAN technique has been applied in a prediction study of success in graduate education. In a study by Houston (10) two variations of JAN were investigated--Normative JAN and Ipsative JAN. The purpose of the Normative JAN study was to determine the extent to which a policy regarding graduate admission standards existed among selected graduate faculty members at Colorado State College (now University of Northern Colorado). Basically, three sets of independent profile variables were used: (1) biographical data, (2) test data, and (3) major subject field data. Results from the Normative JAN study indicated essentially one policy was present in the group of judges.

The Ipsative JAN study used for its dependent variable the rankings submitted by the judges who were requested to rank, without access to the three sets of independent profile variables used in the Normative JAN study, the doctoral graduates on the basis of a personal knowledge. It was the intent in this phase that the ratings or rankings be loaded with personality factors not readily available in the Normative JAN study. Results of this phase were statistically significant, though weak from a predictive standpoint. The practical significance of the Ipsative JAN study was in the suggestion of new directions for subsequent research.

Williams, Gab, and Linden (23) replicated Houston's Normative study at the University of North Dakota and sought to determine the policy of a university doctoral admissions board. Twelve members of the graduate faculty evaluated each graduate student's profile and place in into one of seven criterion categories (Q-Sort). Each rater's policy was assessed or captured and the raters were grouped into appropriate clusters by the JAN process. The investigators found that at least two separate judgmental systems were present.

A further illustration of the versatility of the technique is provided in a study by Stock (18) who sought to determine if systematic differences existed in the placement policies for special education students among special education personnel (teachers, administrators, and the members of the special education screening committee) responsible for placing the students in the public schools of Cheyenne, Wyoming. Coltvet (3) used JAN techniques in the identification and analysis of the consultant ratings of elementary student teachers at the University of Northern Colorado. The policies involved in the selection of faculty members at two small private colleges were investigated by Heim (13). Using JAN procedures, Chang (2) designed a study to determine whether individuals serving in different official capacities in the State of Colorado had differing attitudes toward selection criteria for awarding college financial grants. Keelan et al. (16) captured the leadership policies of selected firemen in the State of Colorado with the use of JAN.

The question of what is pornographic is indeed a difficult policy area and it was investigated by Houston and Houston (8) who used JAN as a methodology by testing this technique with three groups concerned with this issue. These groups included doctoral students majoring in Psychology, Counseling and Guidance at the University of Northern Colorado, lawyers and police officers from the city of Greeley, Colorado. The JAN technique proved to be surprisingly effective in capturing and clustering the policies (specific and complex) of the judges from the three groups identified. As expected, many policies were present.

The policy-capturing methodology has been extended to cover the situation in which raters are required to make multidimensional rating responses, i. e., they must rank or rate each profile on more than one output dimension. Johnson and King (15) developed as part of a team dissertation on multidimensional policy-capturing a procedure which groups multidimensional policies expressed by k-judges. The technique was identified as Canonical Judgment Analysis (C-JAN) and it represents an extension of Judgment Analysis procedures to two or more dimensions.

2. Judgment Analysis--Evaluation Applications

The problem of evaluating curriculum packages was explored by Torgunrud (19) in a doctoral dissertation completed at the University of Californiaat Los Angeles under the direction of Dr. John I. Goodlad. Torgunrud identified from the educational literature the following independent variables as important dimensions of any curriculum package or set of materials which are under consideration for possible adoption. These include: (1) valid and significant content, (2) significant elements of organization, (3) sequence providing a cumulative effect, (4) integration providing horizontal relationships, (5) value position clearly stated, (6) specificity providing direction, (7) flexibility providing alternatives, (8) accommodation of student differences, (9) accommodation of teacher competencies, (10) accommodation for student participation, and (11) provision for measurement of achievement. After defining the variables, Torgunrud generated a sample of 100 profiles, each described on the 11 variables, by using techniques described by Naylor and Wherry (17) for simulating stimuli with specified factor structure. An example of such a profile appears below:

SAMPLE CURRICULUM PACKAGE

Descriptors	Weak Average Strong 1 2 3 4 5 6 7 8 9 10
1. Valid and significant content	...4
2. Significant elements of organization	5
3. Sequence providing a cumulative effect	7
4. Integration providing horizontal relationships	6
5. Value position clearly stated	10
6. Specificity providing direction	7
7. Flexibility providing alternatives	5
8. Accommodation of student differences	..3
9. Accommodation of teacher competencies	6
10. Accommodation of student participation	9
11. Provision for measure of achievement	...4

SAMPLE CURRICULUM PACKAGE (Continued)

Judge's Overall Rating of Curriculum Package (Circle one)	Weak Average Strong 1 2 3 4 5 6 7 8 9 10

Each of the participating judges (school superintendents) was requested to rate each of the sample curriculum packages on a ten-point scale. The JAN procedures enabled the investigator to determine the relationships of the set of independent variables to the judge's specific policy relative to curriculum packages. In addition, a determination was made how many different evaluation policies were actually present, who were the members expressing each of the policies, and what were the areas of agreement and disagreement in the different policies.

In another evaluation at the University of California at Los Angeles, Duff (7) utilized JAN techniques to capture both the teacher-hiring policies (Ex Ante) of selected administrators and the administrators' evaluation policies (Ex Post) of teachers' on-the-job performance after their first year of paid teaching experience. Both types of policies (hiring and job performance) were analyzed for elements of predictive validity by the investigator.

The concept of a group or collective judge was introduced in a study by Houston and Gilpin (13) and in a follow-up study by Houston et al. (12) in which the investigators sought to capture the evaluation policies of selected groups of students who were rating faculty members at the University of Northern Colorado. Rather than use each student as a separate judge, students were classified by their grade level and by the school or college in which they enrolled as a major. In both studies, regardless of the grouping of the students to form a collective judge, it was found that only one policy was present.

Teacher effectiveness policies of faculty members were captured in a study by Houston (9). Profiles of sixty hypothetical faculty members were generated in which each faculty member was described in terms of teaching skills, student advising skills, institutional service, and professional activities. Three policies were found to exist in the faculty members who participated in the study.

Houston and Bentzen (11) in a study at the University of California at Los Angeles used JAN techniques to capture the different rating policies present in a team of evaluators assigned to evaluating the teaching effectiveness in cluturally disadvantaged junior high mathematics classes. Another evaluation study involving JAN occurred in the Charles F. Kettering Foundation Project (League of Cooperating Schools) at the University of California at Los Angeles. The policies of the four evaluators were captured by Houston et al. (14) as they rated each of eighteen schools on several independent variables and one dependent variable of overall effectiveness.

The effectiveness of JAN in capturing and clustering raters' policies was investigated by Dudycha (6) in a Monte Carlo evaluation of JAN as a methodology. Dudycha's outcomes show that the grouping process begins to break down when there are fewer than 200 stimuli being evaluated or 100 if ten or more stimulus dimensions are used. Consequently, the researcher using JAN must be concerned with the number of stimulus dimensions used in relationship to the stimuli being evaluation. It is the present recommendation of the writers that a minimum of 100 stimuli be available for each judge on a maximum of 10 stimulus dimensions.

3. Evaluation of Research Proposals and Judgment Analysis

A very basic question in the evaluation of research proposals is to determine which proposals should get supported. In answering this question a typical starting point is often to ask the team of evaluators to rate each proposal on a set of characteristics. An example of such a set of characteristics for proposals at the National Cancer Institute on which a team of evaluators would rate, was presented by Dr. John C. Bailar, III, Deputy Director for Cancer Control, National Cancer Institute (1). The dimensions or characteristics identified include: (1) need, (2) relevance, (3) feasibility, (4) number benefitted, (5) how much they would benefit, (6) acceptability, (7) potential, (8) cost benefit, (9) ease of evaluation, and (10) originality. A question often left unanswered (or answered inadequately) when data are available which are descriptive of a specific proposal, is how are each of the ten dimensions to be weighted in making a decision to fund or not to fund the project. JAN seems especially equipped as a technique to not only seek out initially the policies which may be present among the officials making decisions about funding but also to specify ultimately a single policy reflecting the priorities or collective attitude of the group.

A strategy for utilizing JAN in this situation would be to start with the set of ten characteristics and generate a simulated sample of approximately 200 representative proposals possessing a required factor structure. [See the dissertation of Torgunrud (19) and the article of Naylor and Wherry (17) for a description of the procedure required to generate simulated stimuli. However, other approaches are available depending on the availability of stimuli and the requirements of the factor structure.] Have each of the officials involved in the decision-making process rate each of the 200 proposals. [An extremely difficult assignment for the official (who serves as judge or rater) would be the assignment that he rank order the complete set of 200. A more reasonable and practical request would be for each judge to place each of the 200 proposals into different ordered categories (usually a minimum of five). While original studies involving JAN required each judge to rank order the complete set of profiles or proposals, experience with the technique has suggested that JAN is still quite effective when the assignment involves the placing of profiles into ordered categories, even as few as five categories. Judges find it too difficult to rank order a large sample of stimuli and they typically encounter little difficulty in making an assignment involving a few ordered categories.]

An example of a sample proposal or profile described on the ten characteristics appears below:

SAMPLE RESEARCH PROPOSAL

Descriptors	Weak Average Strong 1 2 3 4 5 6 7 8 9 10
1. Need	7
2. Relevance	6
3. Feasibility	3
4. Number benefitted	8
5. How much they would benefit	4
6. Acceptability	5
7. Potential	9
8. Cost benefit	4
9. Ease of Evaluation	7
10. Originality	..2
Judge's Overall Rating of Proposal (Circle only one)	Weak Average Strong 1 2 3 4 5 6 7 8 9 10

After each of the officials serving as judges has completed the rating assignment for the 200 simulated stimuli, begin the JAN analysis. [For a description of the computer program for JAN, see Vlahos and Houston (20).] From the JAN printout and using the recommendations by Ward and Hook (22) for determining the number of policy groupings, identify the number of captured policies and the judges associated with each of the policies. Using the knowledge of how many policies are present and who are the judges in the different policy groupings, the investigator should then attempt to explain each of the captured policies. A variety of approaches have been used to explain the different captured policies. These approaches, depending on the research question and the factor structure relating the set of independent variables, include path analysis (25), principal components analysis (4), stepwise regression analysis (5), setwise regression analysis (24), subjective hierarchical analysis (13), all possible regressions (5), the backward elimination or unique contribution analysis (5, 21), the forward selection (5), and the stagewise regression procedure (5).

Return to the judges and explain the different policies with particular emphasis on areas of agreement and disagreement. At this point identify specifically the profiles whose varied rankings have resulted in the separate policies which were captured. Ask the group to arbitrate the rankings where differences have occurred, and not the weightings to be applied to the independent variables. After the group agrees on the rankings for the simulated proposals, complete a new regression analysis by using the new rankings as the dependent variable and thereby obtain the proper weights associated with the joint policy.

4. Instructional Program Evaluation and Judgment Analysis

Another potential application of JAN techniques is in the area of program evaluation, especially in the case involving the use of evaluation experts rendering professional judgments. Two research studies are described by Houston (11) and Houston et al. (14) in which JAN was used as part of the summative evaluation of the projects. Both programs relied on the expertise of educational evaluation experts who were required to make ratings on several dimensions or variables describing the school or classroom program. These variables were later used as independent predictors of the dependent variable (which was an overall rating of the school or classroom program) in the policy statement of each judge which was captured with the assistance of JAN. Thus with the aid of JAN procedures the investigators were not only able to capture the evaluation policies present in the team of evaluation experts, but also were able to determine which of the independent variables were making significant contributions to the overall success of the projects. Instead of providing much input and output information separately, JAN was able to tie together and relate those input aspects of the program responsible for the ultimate impact of the program. Indeed, JAN can and does provide another dimension to program evaluation and is recommended for consideration in programs which rely heavily on professional judgments of experts in the final evaluation.

5. Summary

Judgment Analysis (JAN) was described as a vehicle for capturing policy (ies) in a group of decision-makers. Several studies were presented which stronly support consideration of its use in the area of proposal and program evaluation. Examples how the technique might be applied to proposal and instructional program evaluation were given.

References

1. Bailar, Dr. John C., III. Federal Cancer Control Program. Seminar Presentation, Department of Epidemiology and Public Health, Yale University, February 18, 1974.
2. Chang, N-K. Hierarchical groupings of judges according to selected criteria for financial aid awards. Unpublished doctoral dissertation, University of Northern Colorado, 1970.
3. Coltvet, A. J. The judgment analysis regression technique applied to evaluation variables for student teachers. Unpublished doctoral dissertation, University of Northern Colorado, 1970.
4. Connett, W. E., Houston, S. R., & Shaw, D. G. The use of factor regression in data analysis. Multiple Linear Regression Viewpoints, 1972, 2 (3), 46-49.
5. Draper, N. R., & Smith, H. Applied regression analysis. New York: John Wiley & Sons, Inc., 1966, 164-177.
6. Dudycha, A. L. A monte carlo evaluation of jan: a technique for capturing and clustering raters' policies. Organizational Behavior and Human Performance, 1970, 5, 501-506.
7. Duff, W. L. A quantitative study of teacher selection and evaluation policies at a suburban elementary school district. Unpublished doctoral dissertation, University of Callifornia at Los Angeles, 1969.
8. Houston, J.A., & Houston, S. R. Judgment analysis and pornography. Multiple Linear Regression Viewpoints, 1973, 4 (3), 18-31.
9. Houston, S. R. Identifying faculty policies of teaching effectiveness. MLR Viewpoints, 1973, 4 (3), 32-38.
10. ___________. The judgment analysis technique applied to the admission variables for doctoral students at Colorado State College, 1963-1966. Unpublished doctoral dissertation, Colorado State College, 1967.
11. Houston, S. R., & Bentzen, M. M. Teaching effectiveness in culturally deprived junior high mathematics classes. Journal of Experimental Education, 1969, 38, 73-78.

12. Houston, S. R., Crosswhite, C. E., & King, R. S. The use of judgment analysis in capturing student policies of rated teacher effectiveness. Journal of Experimental Education, in press.

13. Houston, S. R., & Gilpin, J. W. Hierarchical groupings of students according to their policy of rated teacher effectiveness. SPATE Journal, 1971, 10(2), 38-53.

14. Houston, S. R., Roy, M. R., & Duff, W. L. Judgment analysis as a technique for evaluating school effectiveness. Journal of Experimental Education, 1972, 40, 56-61.

15. Johnson, J. W., & King, R. S. Multiple criterion judgment analysis for the educational researcher. Unpublished team doctoral dissertation, University of Nothern Colorado, 1973.

16. Keelan, J. A., Houston, T. R., & Houston, S. R. Leadership policies as perceived by firemen. Colorado Journal of Educational Research, 1973, 12, 20-23.

17. Naylor, J. C., & Wherry, R. L. The use of simulated stimuli and the JAN technique to capture and cluster the policies of raters. Educational and Psychological Measurement, 1965, 25, 969-986.

18. Stock, G. C. Judgment analysis for the educational researcher. Unpublished doctoral dissertation, Colorado State College, 1969.

19. Torgunrud, E. A. Criteria guiding curriculum decisions of selected school superintendents. Unpublished doctoral dissertation, University of Calfiornia at Los Angeles, 1971.

20. Vlahos, G. E., & Houston, S. R. Computer programs for type a and type b judgment analysis. Unpublished manuscript.

21. Ward, J. H., Jr. Multiple linear regression models. In H. Borko (ed.) Computer applications in the behavioral sciences. Englewood Cliffs, N. J.: Prentice-Hall, 1962, 204-237.

22. Ward, J. H., Jr., & Hook, M. E. Application of an hierarchical grouping procedure to a problem of grouping profiles. Educational and Psychological Measurement, 1963, 23, 69-81.

23. Williams, J. D., Gab, D., Linden, A. Judgment analysis for assessing doctoral admission policies. Journal of Experimental Education, 1969, 38(2), 92-96.

24. Williams, J. D., & Linden, A. Regression computer programs for setwise regression, analysis of covariance with multiple covariates, and two-way analysis of variance and covariance with disproportionate cell frequencies. Paper presented, AERA Association, 1974.

25. Wright, S. Path coefficients and path regressions: alternative or complementary concepts. Biometrics, 1960, 16, 189-202.

Multiple Linear Regression Viewpoints
Vol. 5, No. 3, 1974

UNIQUE VARIANCES OF CREATIVITY AND DOGMATISM FOR PREDICTING COUNSELING SUCCESS

Joseph F. Maola
Assistant Professor of Education
Duquesne University
Pittsburgh, Pennsylvania

by

David M. Weis
Associate Professor of Education
The University of Akron
Akron, Ohio

Numerous research attempts by counselor educators and counseling psychologists have attempted to determine the predictors of counseling success. The limitations for these previous research attempts are readily listed by the researchers and are usually: (1) limitations of assessment instruments and (2) difficulty for determining counseling success. Each researcher has also carefully limited the generalizability of the study. There has usually been one common statistical method of accounting for predictor variance; stepwise multiple regression. While researchers readily list their limitations of design, instrumentation and generalizability the presenters for this paper have not noted any statement of the limitations by the researchers of stepwise multiple regression.

Because of the nature of stepwise multiple regression, which maximizes the variance accounted for by the order of variable placement, one usually finds that the variance accounted for (R^2) is an overestimate. Only if the variables are mutually exclusive of one another can the unique variances of the variables be accounted for by stepwise multiple regression. The purposes for this present paper are: (1) compare the results of stepwise multiple regression and stepwise multiple regression of unique variances for predicting counseling success from creativity and dogmatism, and (2) to describe the unique prediction of creativity and dogmatism for predicting counseling success. Stepwise multiple regression demonstrates to the reader the predictor variances of the variables as they interact with one another; stepwise multiple regression of unique variances describes the predictor variance as these variables are distinct from one another.

Method

Twenty-three graduate students seeking master's degrees in guidance and counseling were selected as subjects for the research. Each subject was administered the Torrance Test of Creative Thinking figural B (1966) and the Rokeach D-scale (1960). Every test was scored by a graduate student not involved in the research. The Torrance figural B was scored according to its four areas: Fluency, Flexibility, Originality and Elaboration. Each area of the Torrance figural B and a total score for the Rokeach D-scale were used as the five (5) independent predictor variables. (The reporters note the possible existence of limitations of instrumentation.)

The dependent variable for the research was counselor success. Each student was ranked by a professor and three doctoral students (interobserver reliability = .92) as most successful to least successful. The subjects' scores on the five independent variables were also rank-ordered for each student. Stepwise Multiple Regression was then calculated to determine the prediction potentials of the independent variables. The regression equations were:

Model 1 $Y_1 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + E$

Model 2 $Y_1 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + E$

Model 3 $Y_1 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + E$

Model 4 $Y_1 = a_0U + a_1X_1 + a_2X_2 + E$

Model 5 $Y_1 = a_0U + a_1X_1 + E$

Where Y_1 = rank order score for counseling success

X_1 = rank order fluency score

X_2 = rank order flexibility score

X_3 = rank order originality score

X_4 = rank order elaboration score

X_5 = rank order dogmatism score

The variables were also calculated by using stepwise multiple regression of unique variance. The equations were: Model 1 - $Y_1 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7 + \ldots + a_{31}X_{31}$ followed by stepwise deletion of variables X_5, X_4, X_3, X_2, X_1 so that the final equation was: Model 6 - $Y_1 = a_0U + a_6X_6 + a_7X_7 + \ldots + a_{31}X_{31}$.

Where:

$X_6 = X_1 * X_2$

$X_7 = X_1 * X_3$

$X_8 = X_1 * X_4$

$X_9 = X_1 * X_5$

$X_{10} = X_2 * X_3$

$X_{11} = X_2 * X_4$

$X_{12} = X_2 * X_5$

$X_{13} = X_3 * X_4$

$X_{14} = X_3 * X_5$

$X_{15} = X_4 * X_5$

$X_{16} = X_1 * X_2 * X_3$

$X_{17} = X_1 * X_2 * X_4$

$X_{18} = X_1 * X_2 * X_5$

$X_{19} = X_1 * X_3 * X_4$

$X_{20} = X_1 * X_3 * X_5$

$X_{21} = X_1 * X_4 * X_5$

$X_{22} = X_2 * X_3 * X_4$

$X_{23} = X_2 * X_3 * X_5$

$X_{24} = X_2 * X_4 * X_5$

$X_{25} = X_3 * X_4 * X_5$

$X_{26} = X_1 * X_2 * X_3 * X_4$

$X_{27} = X_1 * X_2 * X_3 * X_5$

$X_{28} = X_1 * X_3 * X_4 * X_5$

$X_{29} = X_1 * X_2 * X_4 * X_5$

$X_{30} = X_2 * X_3 * X_4 * X_5$

$X_{31} = X_1 * X_2 * X_3 * X_4 * X_5$

The results of the analyses are presented in the next section. (The investigators wish to state the possible limitations of ranking the counselors and that generalizability is limited to structured counselors rated by structured counselors.)

Results

The original use of stepwise multiple regression presented the investigators the following results: (1) dogmatism accounted for 29 percent of the variance, and (2) creativity (taken in general) accounted for 11.1 percent of the variance (see Table I). These results are interactive results and should be generalized for predicting the variance of dogmatism and the four aspects of creativity as they interact overall.

Table I

Stepwise Multiple Regression Analysis for Predicting Counseling Success from the Independent Variables

Step	Predictor Variables	R^2	Decrease in R^2
1	(Full Model) Fluency, Flexibility, Originality, Elaboration, Dogmatism	.401	-
	Restricted Models		
2	Fluency, Flexibility, Originality, Elaboration	.111	.290
3	Fluency, Flexibility, Originality	.031	.080
4	Fluency, Flexibility	.006	.025
5	Fluency	.005	.001
6	-	-	.005

The results of the analysis of unique variance provided these results: (1) dogmatism accounted for 2.6 percent of the variance, and (2) creativity accounted for 3.7 percent of the variance (see Table II).

Table II

Stepwise Multiple Regression Analysis of Unique Variance For Predicting Counseling Success

Step	Predictor Variables	R^2	Decrease in R^2
1	(Full Model) Fluency, Flexibility, Originality, Elaboration, Dogmatism, All Variable Interactions	.880	-
	Restricted Models		
2	Fluency, Flexibility, Originality, Elaboration, All Variable Interactions	.854	.026
3	Fluency, Flexibility, Originality, All Variable Interactions	.853	.001
4	Fluency, Flexibility, All Variable Interactions	.842	.011
5	Fluency, All Variable Interaction	.837	.005
6	All Variable Interactions	.817	.020

Discussion

The results indicate that dogmatism and creativity do account for some of the variance for counseling success. Stepwise analyses can determine both the overall effect of the two variables for predicting success (stepwise) and the unique effect of the two variables (stepwise unique).

There have been many studies to determine the predictors for counseling success. The investigators for this present paper feel that not only overall effects are important but so are the unique effects and that both should be reported in all of these studies.

REFERENCES

Rokeach, M. The open and closed mind. New York: Basic Books Inc., 1960

Torrance, E.P. Torrance tests of creative thinking, figural test booklet B. Princeton, N.J.: Personnel Press, Inc., 1966.

Multiple Linear Regression Viewpoints
Vol. 5, No. 3, 1974

PART, PARTIAL, AND MULTIPLE CORRELATION IN COMMONALITY ANALYSIS OF MULTIPLE LINEAR REGRESSION MODELS

Samuel R. Houston
University of Northern Colorado

and

James T. Bolding, Jr.
University of Arkansas

ABSTRACT

The method of commonality analysis which partitions variance of a linear model is presented as a technique for explaining and analyzing data. An example with hypothetical data is presented and analyzed. The interrelationships and roles of part, partial and multiple correlation to the process of commonality analysis are identified.

In their search for explanation of phenomena behavioral scientists have attempted to determine the relative importance of explanatory variables. A variety of strategies have been suggested and various criteria for the importance of variables have been used. [See Darlington (1968) for a review and discussion of attempts to determine the relative importance of variables.] One approach, called commonality analysis (CA), which was developed by Mood (6), was used extensively in the Coleman Report (1). The approach uses the ideas of unique contribution of an independent variable and the common contribution of a subset of variables.

Unique Contribution and Part Correlation

Given a multiple linear regression (MLR) model, the unique contribution of an independent variable can be defined as the difference between two squared multiple correlation coefficients, the first squared multiple correlation coefficient involves all the independent variables in the system and the second squared multiple correlation coefficient is calculated with all the independent variables except the one whose unique contribution is being sought. [See Ward (7) for a discussion of unique contribution of an independent variable.] As the variance attributed to it when it is entered last in the regression equation, the unique contribution is actually a squared semipartial correlation (or a squared part correlation coefficient) between the dependent variable and the variable of interest, after partialing all the other independent variables from it. Thus, this squared semipartial correlation measures the increase in the squared multiple correlation that is achieved by adding the variable X_i to the set of independent variables.

Consider the MLR model in which $Y = f(X_1,X_2,X_3)$. Let $R^2(X_1,X_2,\ldots,X_k)$

designate the squared multiple correlation of Y with variables $X_1, X_2, \ldots, X_k$. For example, $R^2(X_1,X_2,X_3)$ is the squared multiple correlation of Y with variables X_1, X_2, and X_3. Let $U(X_i)$ be defined as the unique contribution of variable X_i. With the case of three independent variables, the unique contribution of X_3 is defined as follows:

$$U(X_3) = R^2(X_1,X_2,X_3) - R^2(X_1,X_2)$$

Similarly, the unique contribution of X_2 is defined as follows:

$$U(X_2) = R^2(X_1,X_2,X_3) - R^2(X_1,X_3)$$

Common Contribution and Commonality Analysis

The concept of unique contribution of a variable and the common contribution of a set of independent variables are tied together in CA. Unique contribution as it is used in CA will always involve a single variable although unique contribution could be extended to two or more variables. On the other hand, the concept of the common contribution will always involve two or more independent variables. Let $C(X_1,X_2)$ be defined as the second-order commonality of variables X_1 and X_2. $C(X_1,X_2,X_3)$ is an example of a third-order commonality. Armed with some set theory, one can determine that, in general, the number of components required for CA is equal to $2^k - 1$, where k is the number of independent variables. [See Fruend (4) for a discussion of sets and the formula for determining the number of proper subsets.] Thus, with a MLR model involving three independent variables, there are $2^3 - 1 = 7$ components, three of which are unique, three are second-order, and one is third order. [Generalized formulae for determining commonalities of any order are developed by Mood (6) and are not presented here.] The elegance of CA is that it enables the investigator to express the squared multiple correlation coefficient for any MLR model as a sum of unique and common contributions which are all mutually exclusive and collectively exhaustive. For the three independent variable model, the following relationship can be established:

$$R^2(X_1,X_2,X_3) = U(X_1) + U(X_2) + U(X_3) + C(X_1,X_2) + C(X_1,X_3) + C(X_2,X_3) + C(X_1,X_2,X_3)$$ where

$$C(X_1,X_2,X_3) = R^2(X_1) + R^2(X_2) + R^2(X_3) - R^2(X_1,X_2) - R^2(X_1,X_3) - R^2(X_2,X_3) + R^2(X_1,X_2,X_3)$$ and

$$C(X_1,X_2) = R^2(X_1,X_3) + R^2(X_2,X_3) - R^2(X_3) - R^2(X_1,X_2,X_3)$$ and

$$C(X_1,X_3) = R^2(X_1,X_2) + R^2(X_2,X_3) - R^2(X_2) - R^2(X_1,X_2,X_3)$$ and

$$C(X_2,X_3) = R^2(X_1,X_2) + R^2(X_1,X_3) - R^2(X_1) - R^2(X_1,X_2,X_3)$$

In general, it can be shown that any unique contribution or common contribution is merely a function of the squared multiple correlation coefficients involving the appropriate independent variable(s).

Part and Partial Correlation Coefficients

Coefficients of partial correlation measure the correlation between the dependent variable and each of the several independent variables, while eliminating any linear tendency of the remaining independent factors to obscure the relation. Partial correlation attempts to measure the importance of each of the several variables be determining how much it reduces the variation after all the other variables except it are taken into account. Let the square of the partial correlation between X_3 and Y after X_1 and X_2 have been taken into account, be designated as $P^2(X_3)$. What is the relationship between this partial correlation squared and $U(X_3)$? Its relationship to the unique contribution of X_3 (or squared part correlation) can be shown to be as follows:

$U(X_3) = R^2(X_1,X_2,X_3) - R^2(X_1,X_2)$ whereas

$$P^2(X_3) = [R^2(X_1,X_2,X_3) - R^2(X_1,X_2)]/[1 - R^2(X_1,X_2)]$$
$$= U(X_3)/[1 - R^2(X_1,X_2)]$$

Thus, $U(X_3) = P^2(X_3).\ [1 - R^2(X_1,X_2)]$

[See Ezekiel and Fox (3) and Winer(8) for details.] A comparison of the unique contributions with the corresponding square of the partial correlation coefficients shows that the square of the partial correlation coefficients will always be equal to or greater than the unique contributions as the unique contributions are divided by unexplained variances (less than or equal to 1.0) in order to determine the square of the partial correlation coefficients.

Example with Hypothetical Data

A set of hypothetical data involving three independent variables and one dependent variable appears in Table 1.

TABLE 1
HYPOTHETICAL DATA FOR MLR MODEL WITH THREE INDEPENDENT VARIABLES

Y	X_1	X_2	X_3
2	1	3	3
1	2	1	2
3	.3	2	1
5	4	4	4
4	5	5	5

In Table 2 the matrix of intercorrelations for the data described in Table 1 are presented.

TABLE 2

INTERCORRELATION MATRIX FOR HYPOTHETICAL DATA

	X_1	X_2	X_3	Y
X_1		0.70	0.60	0.80
X_2			0.90	0.80
X_3				0.60

The following are squared multiple correlation coefficients essential for CA. These include: 1) $R^2(X_1,X_2,X_3) = 0.7634$; 2) $R^2(X_1,X_3) = 0.6625$; 3) $R^2(X_1,X_2) = 0.7529$; and 4) $R^2(X_2,X_3) = 0.7158$. In Table 3 a summary of the data is presented for CA.

TABLE 3

SUMMARY TABLE OF HYPOTHETICAL DATA FOR COMMONALITY ANALYSIS

	1	2	3	y
$P^2(X_1)$	0.1686			
$P^2(X_2)$		0.2990		
$P^2(X_3)$			0.0425	
$U(X_1)$	0.0476			0.0476
$U(X_2)$		0.1009		0.1009
$U(X_3)$			0.0105	0.0105
$C(X_1,X_2)$	0.2549	0.2549		0.2549
$C(X_1,X_3)$	0.0653		0.0653	0.0653
$C(X_2,X_3)$		0.0120	0.0120	0.0120
$C(X_1,X_2,X_3)$	0.2722	0.2722	0.2722	0.2722
Total (R^2-values)	0.6400 $[R^2(X_1)]$	0.6400 $[R^2(X_2)]$	0.3600 $[R^2(X_3)]$	0.7634 $[R^2(FM)]$

From Table 3 it can be seen that the overall R^2 for the full linear model $[R^2(X_1,X_2,X_3)$ or $R^2(FM)]$, where $Y = f(X_1,X_2,X_3)$, can be partitioned into mutually exclusive and collectively exhaustive additive components which are unique and common contributors. That is, $R^2(FM) = U(X_1) + U(X_2) + U(X_3) + C(X_1,X_2) + C(X_1,X_3) + C(X_2,X_3) + C(X_1,X_2,X_3)$. In addition, the individual R^2s relating each independent variable with the dependent variable can also be partitioned into mutually exclusive additive components which are unique and common contributors. That is, $R^2(X_1) = U(X_1) + C(X_1,X_2) + C(X_1,X_3) + C(X_1,X_2,X_3)$, $R^2(X_2) = U(X_2) + C(X_1,X_2) + C(X_2,X_3) + C(X_1,X_2,X_3)$ and $R^2(X_3) = U(X_3) + C(X_1,X_3) + C(X_2,X_3) + C(X_1,X_2,X_3)$. Furthermore, from Table 3, it is obvious that variable 2 is making the highest unique contribution while variables 1 and 2 are making the highest second-order common contribution. Finally, it should be observed that the contribution from the third-order commonality was in excess of one-third of the total R^2(FM). The kind of partitioning which is illustrated in Table 3, can be extended to a linear model involving k-independent variables.

Problems Associated with Commonality Analysis

Some questions about the use of CA are raised by Kerlinger and Pedhazur (5). As the unique contribution of a variable is defined as the additional amount in proportion of variance accounted for when it is entered last in the regression, it should be obvious that the uniqueness of variables depends on the relations among the specific set of variables under consideration. Addition or deletion of variables in a MLR model can drastically change the unique contribution of a variable. There are also problems associated with common contributions. While it may be possible to explain a second- or a third-order commonality, it is extremely difficult to explain commonalities of higher orders.

REFERENCES

1. Coleman, J.S., Campbell, E.Q., Hobson, C.J. McPartland, J., Mood, A.M., Weinfeld, F.D., & York, R.L. Equality of educational opportunity. Washington, D.C.: U.S. Dept. of Health, Education, and Welfare, Office of Education, U.S. Govt. Printing Office, 1966.

2. Darlington, R.B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.

3. Ezekiel, M., & Fox, K.A. Methods of correlation and regression analysis. New York: Wiley, 1965.

4. Fruend, J.E. Mathematical statistics. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1971.

5. Kerlinger, F.N. & Pedhazur, E.J. Multiple regression in behavioral research. New York: Holt, Rinehart & Winston, Inc., 1973.

6. Mood, A.M. Macro-analysis of the American educational system. Operations Research, 1969, 17, 770-784.

7. Ward, J.H., Jr. Multiple linear regression models. In H. Borko (ed.), Computer applications in the behavioral sciences. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962.

8. Winer, B.J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Multiple Linear Regression Viewpoints
Vol. 5, No. 3, 1974

SOME THOUGHTS ON CONTINUOUS INTERACTION

Keith McNeil
Educational Monitoring Systems

and

Judy McNeil
High/Scope Educational Research Foundation

Multiple Linear Regression can be represented as the product of the two (linear) continuous variables. If a researcher hypothesized that the combined effect (interaction) of two variables (X_2 and X_3) on the criterion would be that described by Hypothesis A in Table 1, the product term would match the expected criterion outcome. But if the researcher hypothesized that the combined effect of X_2 and X_3 was that described by Hypothesis B in Table 1, the product term would not match the expected criterion outcome. (Inverting the products and using a negative weighting coefficient would not produce the desired results.) Therefore the single product term will not suffice for all interactions, and it becomes apparent how important it is to specify the expected interaction. The remaining of this paper will illustrate several ways of allowing for the expected interaction.

It was proposed in an earlier paper (McNeil & McNeil, 1973) that the term which would allow the interaction to be manifested for Hypothesis B was simply the one variable divided by the other variable, similar to the notion of IQ being "mental age" divided by

Table 1. Several hypothesized interaction, with conceptual and numerical examples.

X_2	X_3	X_2*X_3	Hypothesis A Expectation on the criterion	Hypothesis B Expectation on the criterion	(X_2*X_3)	X_2	X_3
Hi	Hi	Hi	Hi	Med	100	10	10
Hi	Lo	Med	Med	Hi	50	10	5
Lo	Hi	Med	Med	Lo	50	5	10
Lo	LO	Lo	Lo	Med	25	5	5

"chronological age." Table 2 indicates that, with sample data, the ratio of the two linear terms does correspond to the expectation on the criterion. Note that it is also the case that reversing one of the linear terms, and then multiplying the reversed scale by the other linear term also results in the same correspondence to the expectations on the criterion. (Rescaling X_2 instead of X_3 would have resulted in scores just reversed from those in column 5; and the weighting coefficient for the product vector would be

Table 2. Hypothesized interaction, with ratio and product vectors both matching the hypothesized interaction.

X_2	X_3	Hypoth-esis B Expectation on the criterion	X_2/X_3	$X_2*(15-X_3)$	X_2	$(15-X_3)$	X_3
Hi	Hi	Med	1	50	10	5	10
Hi	Lo	Hi	2	100	10	10	5
Lo	Hi	Lo	.5	25	5	5	10
Lo	Lo	Med	1	50	5	10	5

negative, rather than positive as will be the case for column 5 vector. It makes more sense to us, though, to reverse the variable which will result in the product corresponding directly, rather than indirectly, to the expectations.)

Table 2 might lead the reader to believe that the "interaction vectors" in column 5 will always be 50 times larger than in column 3, thus yielding equivalent results. This is true for the values in Table 2, but for any intermediate values this will not be the case. For instance, $X_2 = 7$ and $X_3 = 7$ yields a value of 1 for $[X_2/X_3]$, but a value of 56 for $[X_2*(15-X_3)]$. Indeed, the division of one variable by another does not even allow for linear interaction, but for a particular kind of curvilinear interaction, as depicted in Figure 1.

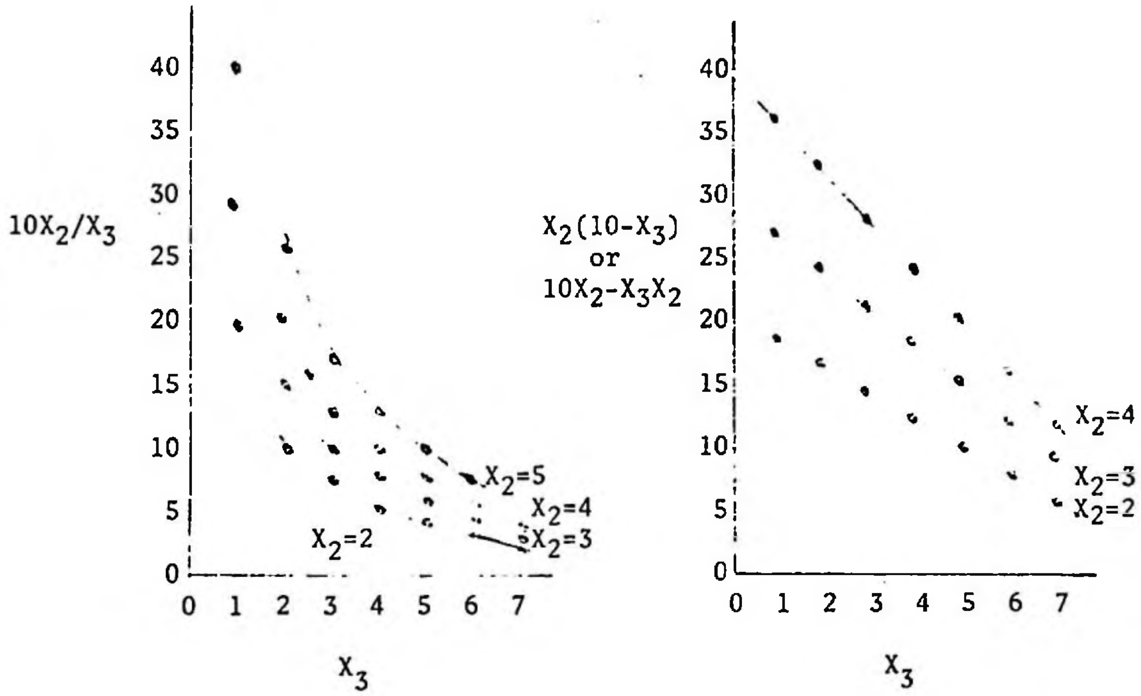

Figure 1. Schematic representations of two kinds of interactions; the ratio of two variables allowing for curvilinear interaction, whereas the product of two vectors allows for linear interaction.

Some fabricated data was analyzed to further establish the point. Before discussing the results, several additional concerns need to be presented. It was not originally clear to us if the choice of the constant used in reversing the one score made any difference. We were further concerned that the place where all lines intersected needed to be represented by a constant in the transformation. These two concerns were empirically examined on the data in Figure 2, with the results appearing in Table 3.

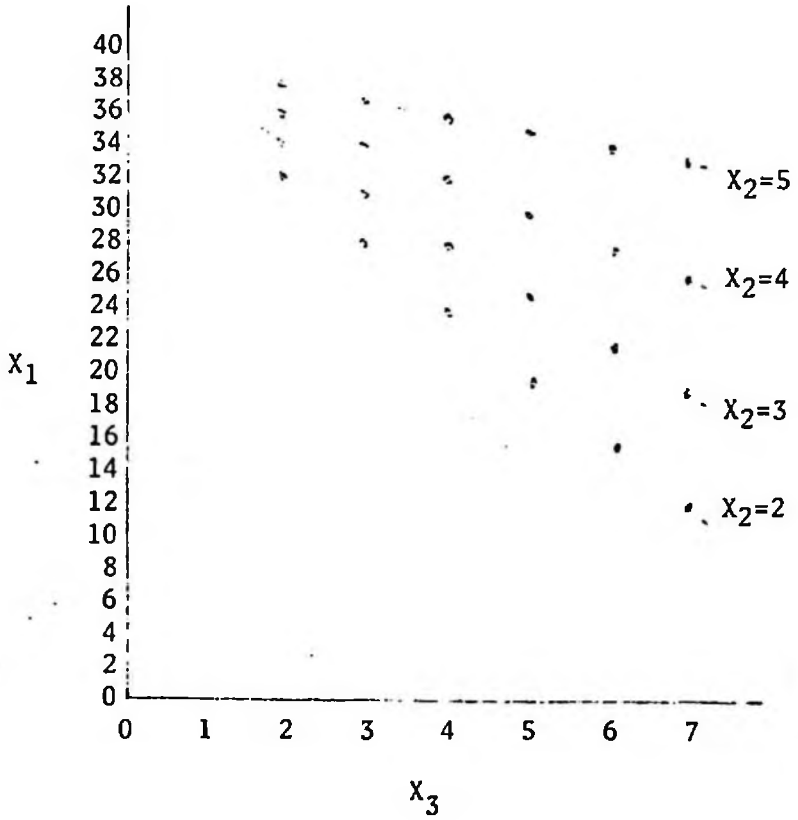

Figure 2. Data which depict a linear interaction, on which models in Table 3 were run.

Table 3. Model, R^2, weighting coefficients, and prediction equations.

Model 1: $R^2 = .95$

$$X_1 = 30.46 + 6.07X_2 - 3.88X_3 - 5.93X_8 \text{ (where } X_8 = X_2/X_3\text{)}$$

Model 2: $R^2 = 1.00$

$$X_1 = 10 + 8X_2 - 6X_3 + 1X_5 \text{ (where } X_5 = 30-(8-X_3)*X_2\text{)}$$

$$X_1 = 40 - 6X_3 + X_2X_3$$

Model 3: $R^2 = 1.00$

$$X_1 = -10 + 6X_2 - 6X_3 + 1X_6 \text{ (where } X_6 = 50-(6-X_3)*X_2\text{)}$$

$$X_1 = 40 - 6X_3 + X_2X_3$$

Model 4: $R^2 = 1.00$

$$X_1 = 0 + 8X_2 - 6X_3 + 1X_7 \text{ (where } X_7 = 40-(8-X_3)*X_2\text{)}$$

$$X_1 = 40 - 6X_3 + X_2X_3$$

Model 1 yields a lower R^2 than the other models because X_8 did not allow for the linear interaction existing in the data. Models 2, 3, and 4 all yielded an R^2 of 1.00 indicating that the concern as to what constant to use in reversing the linear term, and the concern as to what constant to use in allowing all the lines to intersect were of no consequence. The weighting coefficients adjusted to yield the same prediction equation.

Summary

The intent of this paper is not to discredit using the ratio of two variables, but to emphasize again the necessity for specifying one's expectations. To say that interaction is expected is not enough. One must specify if that interaction is linear or curvilinear, and which combinations of the predictor variables will yield high/low criterion scores.

Reference

McNeil, J. T. & McNeil, K.A. A regression analysis of the functional relationship between mother-infant physical contact and infant development. Paper presented to American Psychological Association, Montreal, August, 1973.

Multiple Linear Regression Viewpoints
Vol. 5, No. 3, 1974

STWMULTR: A COMPUTER PROGRAM TO EXPEDITE THE RETRIEVAL OF RESIDUAL SCORES

Carl E. Edeburn and David P. Ochsner
South Dakota State University

Summary - Residual gain analysis was described in general terms and a new computer program, STWMULTR, designed to retrieve and punch residual scores was described. Samples of input and output data cards were included.

One of the most difficult tasks in psycho-educational research investigation has been the measurement of change. When pre and post-testing models have been implemented, the most prevalent application in most analyses has been the use of raw gain scores. Due to the inadequacies of this approach several different solutions, including residual gain analysis, have been proposed.

The residual gain analysis approach has been discussed by Dubois (1957, 1970) and Bakan (1970). Others, (Williams and Maresh, 1972, Buzzahora and Williams, 1973 and Edeburn and Landry, 1974) have applied this technique in test-retest situations using elementary school students' scores as elicited by various cognitive and affective measures.

Essentially, the residual gain method can be concieved as a partial correlation between the group membership variable and the residuals in the posttest data using the pretest as a predictor.

One of the physical limitations of the residual gain application has been the amount of clerical time spent in extracting the residual scores for each student from the computer printout, and repunching them on appropriately identified cards. The present effort was aimed at overcoming this limitation.

To a common version of a stepwise multiple linear regression program (STWMULT) originally adapted from the Scientific Subroutine Package (IBM, 1972), the present authors have added an optional feature which stores and then punches the residual scores for each subject on a new data card. This new program is identified in the S.D.S.U. Computing Center as STWMULTR. The only restriction in the STWMULTR version is that the first card for each subject in the original data set must include 1-8 columns of Alphanumeric I.D. (see Figure 1.).

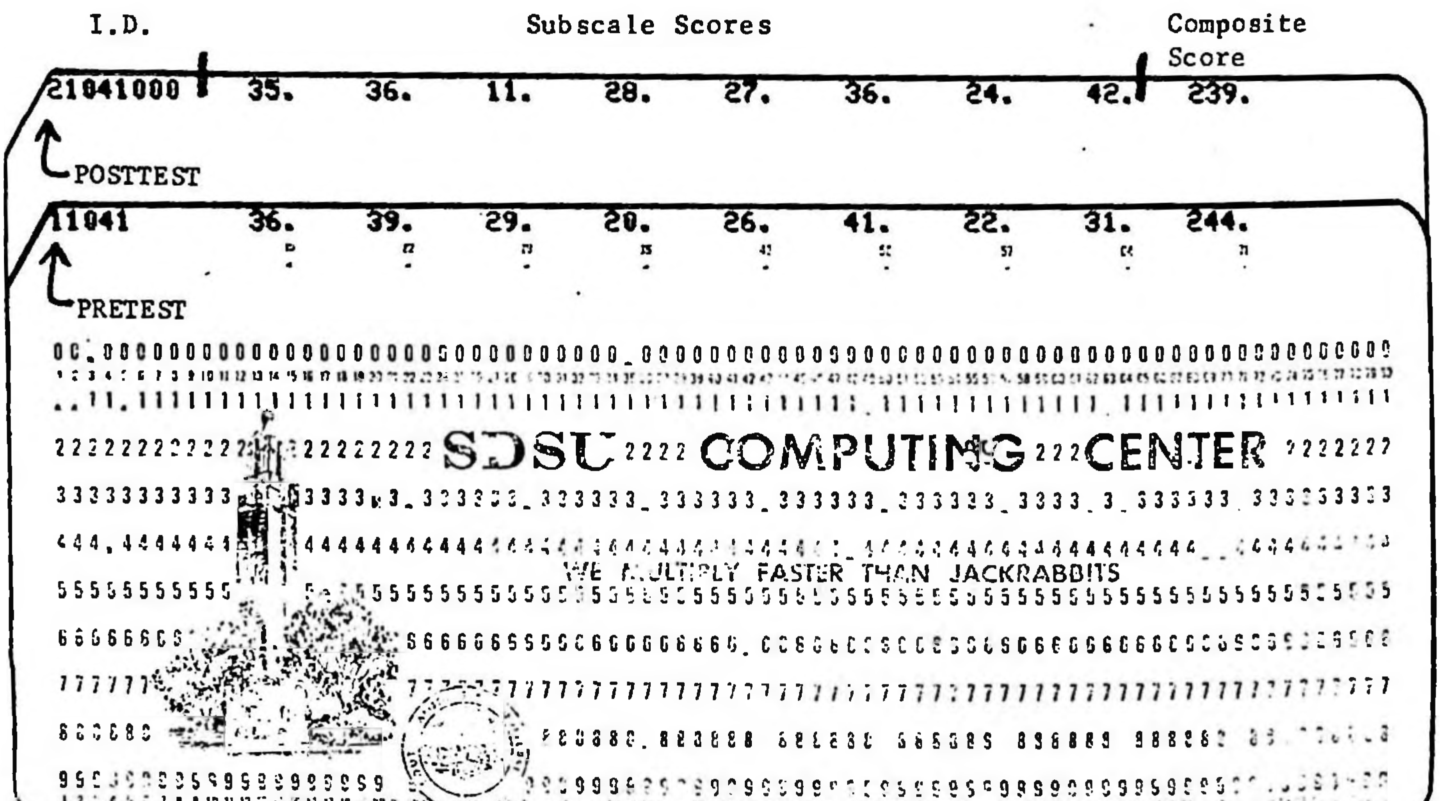

Fig. 1.-- An example of the pre-posttest data for student #104.

Summary

The intent of this paper is not to discredit using the ratio of two variables, but to emphasize again the necessity for specifying one's expectations. To say that interaction is expected is not enough. One must specify if that interaction is linear or curvilinear, and which combinations of the predictor variables will yield high/low criterion scores.

Reference

McNeil, J. T. & McNeil, K.A. A regression analysis of the functional relationship between mother-infant physical contact and infant development. Paper presented to American Psychological Association, Montreal, August, 1973.

REFERENCES

Bakan D. Analysis of covariance in training research. In P.H. Dubois and G.D. Mayo (Eds.), Research strategies for evaluation training. Chicago: Rand McNalley, 1970.

Buzzahara R.C. & Williams J.D. An empirical comparison of residual gain analysis and the analysis of covariance. Multiple Linear Regression Viewpoints, 1973, 4, 10-17.

Dubois P.H. Multivariate correlational analysis. New York: Harper, 1957.

Dubois P.H. Correlational analysis in training research. In P.H. Dubois and G.D. Mayo (Eds.), Research strategies for evaluation training. Chicago: Rand McNalley, 1970.

Edeburn C.E. & Landry R.G. Teacher self concept and student self concept. Paper presented at the annual convention of the American Educational Research Association, April 16, 1974.

IBM Concepts and applications of regressions analysis 3rd. Ed., White Plains, New York: Author, 1972.

Williams, J.D. & Maresh, R.T. A comparison of raw gain scores, residual gain scores, and the analysis of covariance with two modes of teaching reading. Multiple Linear Regression Viewpoints 1972, 2, 2-16.

Multiple Linear Regression Viewpoints
Vol. 5, No. 3, 1974

MULTIPLE REGRESSION TECHNIQUES APPLIED TO TEST THE EFFECT OF THREE TYPES OF SPECIAL CLASS PLACEMENT ON THE READING ACHIEVEMENT OF EDUCABLE MENTALLY RETARDED PUPILS

Joseph D. George*
Columbus College
Columbus , Georgia

ABSTRACT

Multiple regression analysis was used to examine the different effects of special class placement on the reading achievement of Educable Mentally Retarded (EMR) pupils. Self-contained classes, selected academic placement programs and learning resource centers were the types of placement studied. A significant difference was found between the reading scores of EMR boys in learning resource centers and EMR girls in the same classes. The boys scored higher. Further, boys scored higher in learning resource centers than they did in any other placement, and girls scored higher in selected academic placement programs than they did in any other placement.

*Data for this study was acquired as a result of this researcher's participation in the ESEA Title III project, number 45-72-207-2. Appreciation is extended to Dr. Thomas L. Noffsinger, director of the project, for granting permission to use the data.

REVIEW OF LITERATURE

The social and academic value of traditional programs for mentally retarded pupils have long been questioned. On one hand, literature supports the notion that the academic value of self-contained special class placement is less than that of regular class placement. On the other hand, research supports the idea that special class placement promotes for the retarded greater social growth than regular class placement.

The question of academic value was addressed by Lilly (1970) when he expressed the belief that traditional self-contained special classes provides inadequate educational services for educable mentally retarded pupils but furnish satisfactory ones for the severely (trainable) mentally retarded. Attending the same point, Johnson (1962) suggested that specialized training of teachers of mentally retarded pupils placed in special classes has not seemed to result in a better education for those pupils. He stressed the need to reconsider self-contained classes as a method of delivering instructional services to the educable mentally retarded. Moreover, Dunn (1971) advanced the notion that since regular education is becoming better able to deal with individual differences among pupils, that another form of special class placement, or perhaps, regular class placement, be considered for educable mentally retarded youngsters.

The social value of special class placement has been supported by Johnson and Kirk (1951). After studying a school system noted for its progressive administrative and teaching methods, they indicated mentally retarded children attending regular classes were segregated from their peers in those classes and that the segregation was a result of interaction with their peers. Additionally reported was that parents felt their retarded children received less opportunity to learn the three R's. Addressing the same issue, Farson (1945) pointed to a decrease in deviant behavior of 29 mentally retarded pupils who were withdrawn from regular classes in Philadelphia's public schools and placed in self-contained special classes. Implied was that special class placement promoted the social growth of retarded children, thus providing increased opportunities for them to advance academically.

It appears that mentally retarded children benefit socially from special class placement but that such placement inhibits academic growth. Therefore, an issue involved in the efficacy of special class placement seems to be: How can the benefits of special class placement and regular class placement be "married" to provide optimum growth opportunities for educable mentally retarded youngsters?

STATEMENT OF THE PROBLEM

This study sought to identify and examine the impact of different types of special class placement of educable mentally retarded children. Because the academic value of special class placement is most in question, and because reading is a subject common to all schools, this work was limited to comparing the reading achievement of educable mentally retarded children attending different types of special classes.

PROCEDURES

For purposes of this study, three types of special class placements were identified. They were (1) Self-Contained Classes; (2) Selected Academic Placement Programs; and (3) EMR-Learning Resource Centers. Self-contained classes were considered to be ones in which educable mentally retarded pupils:

1. Received all of their instruction in the basic subjects like language arts, math, and social studies from an appropriately certificated teacher.
2. Received instruction in subjects like physical education, art, and music from specialists in each of those fields or from an EMR certificated teacher.
3. Remained with other EMR pupils throughout the entire school day, with the exception of recess, lunch periods or study halls.

Selected academic placement programs were identified as ones in which educable mentally retarded children:

1. Received all of their instruction in the basic subjects like language arts, math and social studies from a teacher who has been appropriately certificated to work with such pupils.
2. Received instruction in subjects like physical education, art, and music from spcialists in each of those fields.
3. Remained with other EMR pupils when receiving instruction in the basic subjects like language arts, math and social studies, but are permitted to join non-retarded pupils when

receiving instruction in subjects like physical education, art and music. They also join non-retarded students during recess and lunch.

EMR-learning resource centers were defined as ones in which pupils:

1. Received instruction in the basic subjects like language arts, math, and social studies from teachers who have been appropriately certificated to work with pupils in those subject areas and who work closely with an EMR certificated teacher to develop a course of instruction for the retarded pupil.
2. Received instruction in subjects like physical education, art, and music from specialists in each of those fields.
3. Remained with non-retarded pupils except for short periods of time when they went to a room equipped with specialized materials and/or equipment to receive highly individualized instruction in a subject area in which they exhibit a weakness.

The Experimental Design

The design of this study was a nonequivalent control group design as noted in Campbell and Stanley's (1972) monograph on experimental design.

Groups of educable mentally retarded children examined by this study were placed into special education classrooms as prescribed by standards set forth by the Ohio Department of Ecuation, Division of Special Education. Hence, assignments of subjects into groups could not be random. However, Campbell and Stanley (1972) have indicated that the greater the number of characteristics in which the control and experimental groups are similar, the more the prior assigned groups can be considered to be equivalent. Because a pre-test is required by the nonequivalent control group design, and because statistical techniques could be used to examine equivalency as related to pre-test scores, the groups were assumed equivalent.

Description of Sample

Because EMR pupils attending schools participating in the ESEA Title III project number 45-72-207-2 during the 1972-1973 school year were the subjects of this study, the population sampled was limited to groups of educable mentally retarded children attending 385 public schools in 8 project areas in the state of Ohio (Noffsinger 1972-1973). Further, the population was limited to educable mentally retarded pupils attending special classes which met the state of Ohio's requirements for one or another of the three types of special class placements being studied. Ninety of the 385 schools met those requirements. All of the groups of educable mentally retarded pupils attending those 90 schools were included in this study.

Sources of Data

The Ohio Special Achievement Inventory, a test that measures the achievement of educable mentally retarded pupils, was used to gather data pertaining to the reading achievement of the subjects. The pre-tests were administered during the months of September and

October of 1972, by test administrators who were trained by Title IV personnel. Post-tests were given during April and May of 1973.

The reading test was administered individually to children in the primary and intermediate grades. Pupils in the junior high grades were administered the test individually. Senior high pupils received the test as a group.

The Metropolitan Achievement Test of reading was administered to all of the children in the sample. This testing was done only once and occurred as closely after the pupil received the Ohio Special Achievement Inventory as possible.

Groups failing to meet the following criteria were eliminated from the sample population:

1. O.S.A.I. pre- and post-test score of 75% or more of the pupils in each group.
2. Two or more boys and two or more girls in each group.

Of the 90 groups in the sample, 30 met the requirements stated above. Ten of those were Self-Contained Classes; eight were Selected Academic Placement Programs; and twelve EMR-Learning Resource Centers. Utilizing a table of random numbers, eight Self-Contained classes and eight EMR-Learning Resource Centers were selected. All eight selected academic placement programs were used.

As shown in Table 1, there were 43 boys and 37 girls attending Self-Contained classes; 50 boys and 36 girls attending Selected Academic Placement Programs; and 43 boys and 38 girls participating in Learning Resource Centers. The total number of boys included in this study was 136. The total number of girls was 111. Although data regarding the population base of each school district examined

Table 1
DESCRIPTION OF THE SAMPLE BY TYPE OF PLACEMENT, CLASS SIZE, SEX AND TYPE OF DISTRICT*

1a. Self-Contained Classes

Class Code	A	B	C	D	E	F	G	H	Totals
No. of Boys	4	8	4	9	2	3	9	4	43
No. of Girls	2	5	5	2	6	5	5	7	37
Total SS	6	13	9	11	8	8	14	11	80
Type of District	C+	C+	C-	C+	C-	N	N	N	

1b. Selected Academic Placement Programs

Class Code	A	B	C	D	E	F	G	H	Totals
No. of Boys	4	9	6	5	10	4	4	8	50
No. of Girls	4	5	6	4	5	2	5	5	36
Total SS	8	14	12	9	15	6	9	13	86
Type of District	C-	C+	C+	C-	C-	N	N	C+	

1c. Learning Resource Centers

Class Code	A	B	C	D	E	F	G	H	Totals
No. of Boys	9	3	2	4	7	6	8	4	43
No. of Girls	2	7	7	4	4	5	6	3	38
Total SS	11	10	9	8	11	11	14	7	81
Type of District	C-	N	N	N	C+	C+	C+	C+	

*C+ = School District having a population base of over 100,000
C- = School District having a population base of under 100,000
N = School District having a population base of under 50,000

were available, there was no useable data by which to classify districts into city, suburban, or rural. Hence, the data were not stratified by types of district.

Statistical Treatment of Data

To establish a degree of equivalency between the prior assigned groups, a one way analysis of variance using Multiple Regression Techniques was computed from the reading pre-test scores. The null hypotheses tested, the regression models used to test them, and the findings are reported in Table 2. For a description of the variables see Table 6.

Reliability and validity data for the Ohio Special Achievement Inventory were unavailable. The Kuder-Ricahrdson procedure, Formula 21, as reported by Thorndike and Hagen (1969), was used to compute reliability. Scores for each level of the test were obtained and are reported in Table 3.

Table 3
O.S.A.I. READING RELIABILITY SCORES

Test Level	r_{11}	SD	n	M
Primary	.83	3.11	12	8.84
Intermediate	.59	2.46	12	7.61
Junior High	.88	3.77	12	7.66
Senior High	.63	2.76	12	5.80

Table 2

MODELS, F-RATIOS, AND R^2 FOR PREDICTING PRE-TEST SCORES OF SUBJECTS

MODELS AND EXPLANATIONS	R^2	df	ALPHA	F	P
Null Hypothesis 1: There will be no significant difference between reading achievement pre-test scores, as measured by the O.S.A.I., of EMR pupils attending Self-Contained Classes, Selected Academic Placement Programs, and Learning Resource Centers.		1/45	.05	2.9903*	.0906
Full: $Y_8 = a_0U + a_3X_3 + a_4X_4 + a_5X_5 + E$	.0623				
Restricted: $Y_8 = a_0U + E$	.0000				
Null Hypothesis 2: There will be no significant difference between the reading pre-test scores, as measured by the O.S.A.I., of EMR boys attending Self-Contained Classes, Selected Academic Programs, and Learning Resource Centers and the reading pre-test scores of girls attending the same classes.		2/46	.05	2.0090*	.1457
Full: $Y_8 = a_0U + a_6X_6 + a_7X_7 + E$	.0803				
Restricted: $Y_8 = a_0U + E$	.0000				

*Not significant at the .05 alpha level.
NOTE: See Table 6 for a description of the variables used.

To compute concurrent validity, the Pearson Product Moment Correlation procedure was used (Newman, Frye and Newman 1973). The reading sections of the Ohio Special Achievement Inventory were correlated with the corresponding sections of the Metropolitan Achievement Tests. Validity data were gathered for each level of the reading test and are reported in Table 4.

Table 4
PRODUCT-MOMENT CORRELATION COEFFICIENTS
COMPUTED BETWEEN THE O.S.A.I. READING TEST
AND THE MAT READING TEST

Test Level	r	n
Primary	.64*	36
Intermediate	.68*	44
Junior High	.58*	38
Senior High	.57*	32

*Significant beyond .01 alpha level for a two-tailed test.

Finally, using Multiple Regression Techniques, analysis of covariance was used to test the research hypotheses. The research hypotheses, regression models used to test them, and findings are reported in Table 5.

Because a significant difference between the reading scores of boys and girls attending learning resource centers was noted (Hypothesis 7, Table 5) the mean reading scores of boys and girls attending those centers are reported in Table 7.

The fact that none of the hypotheses designed to test the simi-

Table 5

MODELS, F-RATIOS, AND R^2 FOR PREDICTING READING SCORES OF SUBJECTS

MODELS AND EXPLANATIONS	R^2	df	ALPHA	F	P
Null Hypothesis 3: There will be no significant difference in reading achievement post-test scores as measured by the O.S.A.I. between EMR pupils attending Self-Contained Classes, Selected Academic Placement Programs, and Learning Resource Centers when co-varying O.S.A.I. reading pre-test scores.		2/44	.05	1.0540	.3571
Full: $Y_1 = a_0U + a_3X_3 + a_4X_4 + a_5X_5 + a_8X_8 + E$	.3559				
Restricted: $Y_1 = a_0U + a_8X_8 + E$	.3251				
Null Hypothesis 4: Male EMR pupils attending Self-Contained Classes, Selected Academic Placement Programs, and Learning Resource Centers will not score significantly different on the O.S.A.I. reading post-test than female EMR pupils attending the same classes when co-varying O.S.A.I. reading pre-test scores.		1/45	.05	3.3082	.0756
Full: $Y_1 = a_0U + a_6X_6 + a_7X_7 + a_8X_8 + E$	.3713				
Restricted: $Y_1 = a_0U + a_8X_8 + E$	.3251				

NOTE: See Table 6 for a description of the variables used.

Table 5 (Continued)

MODELS AND EXPLANATIONS	R^2	df	ALPHA	F	P
Null Hypothesis 5: There will be no significant difference between O.S.A.I. reading post-test scores of EMR boys attending Self-Contained Calsses and O.S.A.I. reading post-test scores of EMR girls attending the same classes when co-varying O.S.A.I. reading pre-test scores.		1/13	.05	0.4816	.4999
Full: $Y_1 = a_0U + a_6X_6 + a_7X_7 + a_8X_8 + E$	.0557				
Restricted: $Y_1 = a_0U + a_8X_8 + E$	.0207				
Null Hypothesis 6: There will be no significant difference between O.S.A.I. reading post-test scores of EMR boys attending Selected Academic Placement Programs and O.S.A.I. reading post-test scores of EMR girls attending the same classes when co-varying O.S.A.I. pre-test scores.		1/13	.05	0.0791	.7829
Full: $Y_1 = a_0U + a_6X_6 + a_7X_7 + a_8X_8 + E$	.0928				
Restricted: $Y_1 = a\ U + a_8X_8 + E$	.0873				

NOTE: See Table 6 for a description of the variables used.

Table 5 (Continued)

MODELS AND EXPLANATIONS	R^2	df	ALPHA	F	P
Null Hypothesis 7: There will be no significant difference between O.S.A.I. reading post-test scores of EMR boys attending Learning Resource Centers and O.S.A.I. reading post-test scores of EMR girls attending the same classes when co-varying O.S.A.I. pre-test scores.		1/13	.05	7.5088*	.0169
Full: $Y_1 = a_0U + a_6X_6 + a_7X_7 + a_8X_8 + E$	.8955				
Restricted: $Y_1 = a_0U + a_8X_8 + E$	.8351				
Null Hypothesis 8: The interaction between sex and type of class will not account for a significantly greater amount of EMR student variance on the O.S.A.I. reading post-test scores than will sex and type of class separately when co-varying O.S.A.I. reading pre-test scores.		7/39	.05	0.0939	.9983
Full: $Y_1 = a_0U + a_8X_8 + a_{10}X_{10} + a_{11}X_{11} + a_{12}X_{12} + a_{13}X_{13} + a_{14}X_{14} + a_{15}X_{15} + E$	.4069				
Restricted: $Y_1 = a_0U + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7 + a_8X_8 + E$	.3969				

*Significant at the .05 alpha level.

NOTE: See Table 6 for a description of the variables used.

Table 6

A DESCRIPTION OF THE VARIABLES USED

Where the Full Model is:

$$Y_1 = a_0U + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7 + a_8X_8 + E$$

The Variables are:

Y_1 = A criterion, O.S.A.I. reading post-test scores

a_0 = a_1 through a_{15} = Partial regression weights

U = The unit vector (a "1" for each sample)

a_3 = The Self-Contained Class, 1 if in the class, zero otherwise

a_4 = The Selected Academic Placement Program, 1 if in the program, zero otherwise

a_5 = The EMR Learning Resource Center, 1 if in the center, zero otherwise

a_6 = Sex, 1 if male, zero otherwise

a_7 = Sex, 1 if female, zero otherwise

a_8 = O.S.A.I. reading pre-test scores

E = Error vector, difference between predicted score and actual score

a_{10} = X(10) = X(3) * X(6) Male in Self-Contained classroom

a_{11} = X(11) = X(4) * X(6) Male in Academic Placement Program

a_{12} = X(12) = X(5) * X(6) Male in EMR Learning Resource Center

a_{13} = X(13) = X(3) * X(7) Female in Self-Contained classroom

a_{14} = X(14) = X(4) * X(7) Female in Academic Placement Program

a_{15} = X(15) = X(5) * X(7) Female in EMR Learning Resource Center

Table 7

THE MEAN READING SCORES OF BOYS AND GIRLS
IN LEARNING RESOURCE CENTERS

Table 11a. (Mean Pre-Test Scores)

GROUPS	BOYS	GIRLS
1	3.7	2.0
2	6.3	10.7
3	8.5	9.7
4	9.0	12.0
5	5.5	5.5
6	8.3	7.4
7	8.4	7.3
8	5.5	7.3
Total	55.2	61.9

Table 11b. (Mean Post-Test Scores)

GROUPS	BOYS	GIRLS
1	4.4	3.0
2	7.7	11.4
3	10.0	10.3
4	11.0	12.0
5	5.4	6.0
6	9.2	6.4
7	10.5	7.8
8	7.7	5.3
Total	65.9	62.2

larity of pre-test scores were rejected provides evidence that the reading pre-test scores of children in each of the three types of special class placements being tested were not different enough to be considered sign at = .05. As indicated by Campbell and Stanley (1972) this provides evidence that the Non-equivalent Control Group design used in this study could be considered more likely to have internal validity. The internal validity of the design used is further supported by the notion that all of the EMR pupils were identified and placed into special class programs pursuant to standards set forth by Ohio's Department of Education, Division of Special Education. However, it must be pointed out that with respect to the similarity of reading pre-test scores, the F-ratio calculated for hypotheses 1 and 2 although not statistically significant at the .05 level were close to the .1 level of significance. Moreover, the manner in which standards for placing educable mentally retarded students into special programs are implemented differs from school district to school district.

As shown in Table 3, the reliability scores for the O.S.A.I. reading test ranged from 159 at the intermediate level to .88 at the junior high level. Because the Kuder-Richardson procedure (Formula 21) is a conservative estimate of reliability, it can be assumed that the observed reliabilities of the reading test was lower than the actual reliability (Thorndike and Hagen 1969). Hence, for purposes of the present research, the Ohio Special Achievement Inventory reading test can be considered reliable. However, utilization of the same tests in other situations may result in different outcomes.

The high correlations reported in Table 4, between Ohio Special Achievement Inventory reading scores and MAT reading scores, provide evidence that, for purposes of this research, concurrent validity for the Ohio Special Achievement Inventory test of reading exists.

With respect to reading achievement, only null hypothesis 7 was rejected. Table 7 shows that EMR boys in Learning Resource Centers demonstrated more gain in reading achievement than EMR girls in the same classes. Implied is that placement into Learning Resource Centers has a different effect on the reading achievement of boys than on that of girls. With the exception of hypothesis 4, other F-ratios obtained were far from being significant. The probability of .0756 for the F-ratio of 3.3082 computed for hypothesis 4, though not significant at the .05 level, indicates some difference may exist in reading achievement between EMR pupils in the three types of placements studied. This difference is partly accounted for by the difference shown to exist between the reading achievement of boys and girls in Learning Resource Centers. Another, unidentified factor might also be involved. Because pupils in primary, intermediate, and junior high and senior high grades were not differentiated between it was assumed that they would all exhibit the same pattern of high and low scores in reading achievement. Perhaps this was not the case. It might be that senior high level students performed better than elementary level students and that combining of their scores cancelled out differences which might otherwise have been obtained. However far from being indicated by findings of the present study the failure to reject reading related hypotheses 3, 5, 6, and 8 might also have been due to differences in the performance level of primary, intermediate, junior high, and senior high level EMR pupils. Moreover, failure to detect differences might have been due to Type I error resulting from the low n involved in the study. The possibility that no differences exist should not be excluded from consideration.

CONCLUSIONS

For the population studied, the data gathered support the following conclusions:

1. Participation in Self-Contained classes did not result in higher gain scores in reading than participation in Selected Academic Placement Programs or Learning Resource Centers.
2. Boys in Learning Resource Centers evidenced higher reading gain scores than girls in the same center.
3. Boys in Learning Resource Centers scored higher in reading than boys in Selected Academic Placement Programs and boys in Self-Contained classes.
4. For the population studied, special class placement had a different effect upon boys than upon girls. Were this study repeated and results similar, it could be concluded that sex ought to be considered a factor in placing educable mentally retarded youngsters into special classes.

References

Articles

Dunn, Lloyd M. "Special Education for the Mildly Retarded--Is Much Of It Justifiable?" in Problems and Issues in the Education of Exceptional Children, ed. by Reginald L. Jones (Boston: Houghton Mifflin Company, 1971), pp. 382-389.

Farson, Mable R. "A Program For Low Ability Children in the Regular Grade," American Journal of Mental Deficiency, 50:1 (July, 1945), pp. 107-114.

Johnson, Orville G. "Special Education for the Mentally Handicapped--A Paradox," Exceptional Children, 29 (October, 1962), pp. 62-29.

Johnson, Orville G. and Samuel A. Kirk. "Are Mentally-Handicapped Children Segregated in the Regular Grades?" Journal of Exceptional Children, 17:2 (October, 1950-May, 1951), pp. 65-68 and 87-88.

Lilly, Stephen M. "Special Education: A Teapot in a Tempest," Exceptional Children, 37 (September, 1970), pp. 43-49.

Books

Campbell, Donald J. and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research (Chicago: Rand McNally and Company, 1972)

Thorndike, Robert and Elizabeth Hagen. Measurement and Evaluation in Psychology and Education (New York: John Wiley and Sons, 1969).

EDITORIAL COMMENT

Isadore Newman

Regression Chi-Square: Testing For A Linear Trend In Proportions In A 2 x 2 Contingency Table, by Houston & Bolding is being reprinted in this issue (Vol.5, No. 3), because the pages were out of sequence when it was printed in the last issue (Vol. 5, No.2).

Multiple Linear Regression Viewpoints
Vol. 5, No. 3, 1974

REGRESSION CHI-SQUARE: TESTING FOR A LINEAR TREND IN PROPORTIONS IN A 2 X C CONTINGENCY TABLE

Samuel R. Houston
University of Northern Colorado

and

James T. Bolding, Jr.
University of Arkansas

ABSTRACT

The usual chi-square test for a 2 X C contingency table can fail to produce statistical significance when, in fact, a significant linear trend in proportions is present in the data. The ordinary chi-square test lacks power in the case when the variables can be considered ordered classifications. A regression chi-square test is described and illustrated with hypothetical data in which the usual chi-square test produced non-significant results even though a significant linear trend was present in the data.

Consider the following hypothetical data. A survey was conducted to gather information from voters (registered and non-registered) concerning support for a local school bond. Each voter was asked to check one of five categories ranging from strongly oppose to strongly favor and each response was used as a measure of support for the school bond. The school district believed that a more appropriate campaign could be conducted if systematic differences were detected on the initial survey. A statistican was consulted ex post facto (as usual!) and it was suggested that the data be arranged in a 2 X 5 contingency table and the usual chi-square test be completed.

Chi-Square Analysis of the Data

Table 1 is a summary table of responses from the survey.

TABLE 1

120 VOTERS CLASSIFIED ACCORDING TO REGISTRATION STATUS AND DEGREE OF SUPPORT FOR LOCAL SCHOOL BOND

Voter Registration Status	Attitude Toward Local School Bond					Total
	Strongly Oppose (SO)	Oppose (O)	Neutral (N)	Favor (F)	Strongly Favor (SF)	
Registered	7	15	25	9	4	60
Non-registered	3	11	25	13	8	60
Total	10	26	50	22	12	120

The calculation of the usual chi-square value produces a value of 4.28, which is not significant for 4 degrees of freedom (d. f.). On the basis of this test the researcher might conclude at this point that there is no evidence to suggest that voter registration status is related to attitude of support for local school bond. Yet, a closer look at the data suggests that a linear trend in proportions might be present. A modification of Table 1 to include information which is suggestive of a linear trend appears in Table 2.

TABLE 2

TESTING A LINEAR REGRESSION ON p_i ON SUPPORT SCORE REGRESSION

Voter Registration Status	Attitude Toward Local School Bond					Total
	SO	O	N	F	SF	
Registered (a_i)	7	15	25	9	4	60
Non-registered	3	11	25	13	8	60
Total (n_i)	10	26	50	22	12	120(N)
Registered Voter Rate ($p_i = a_i/n_i$)	0.70	0.56	0.50	0.45	0.33	0.5($\overline{p}$)
Coded Score (X_i)	1	2	3	4	5	

The proportions in the voter registration rate seem to be decreasing with the ordered categories which reflect attitude toward the local school bond issue. (See the row of Registered Voter Rates in Table 2.) The rates start at 0.70 (Strongly Oppose) and decrease sequentially to a low of 0.33 (Strongly Favor). This relationship is not detected in the usual chi-square test as the chi-square test is not sensitive to ordered categories which may be present in the classification variables.

<u>Two Sample t-test Analysis of the Data</u>

One alternative to the usual chi-square test, a t-test for two independent samples, has been suggested by Cochran (2), Armitage (1) and Yates (4). This approach involves attaching a coded score to each class so that an ordered scale is created. To illustrate from the school bond example, let us assign a 1, 2, 3, 4, 5, respectively to Strongly

Oppose, Oppose, Neutral, Favor, and Strongly Favor classes. Having assigned the scores we may think of the school survey data as consisting of two independent samples of 60 and 60 voters, respectively. In Table 3 the data are presented in coded units and frequencies.

TABLE 3

SCHOOL BOND DATA DISPLAYED BY CODED SCORES

Attitude to School Bond	Voter Registration Status	
	Registered	Non-registered
X_i	f_i	f_i
1	7	3
2	15	11
3	25	25
4	9	13
5	4	8
Total	60	60
Mean ($\overline{X}$)	2.8	3.2

On the X scale the average attitude toward the school bond is 2.8 for the registered voters and 3.2 for the non-registered voters. The difference, $\overline{D}$, is -0.4, with a standard error $\pm$ 0.193 (118 d.f.), computed in the usual way. The value of $t = -0.4/0.193 = -2.08$ is significant at 0.05 level. Contrary to the initial chi-square test, this test reveals a significantly greater negative attitude score for registered voters than for non-registered voters.

Snedecor and Cochran (3) state that the assignment of coded scores is appropriate when (i) the phenomenon in question is one that could be measured on a continuous scale, and (ii) the ordered classification can be regarded as a kind of grouping of this continuous scale by a cruder scale that is the best we can do in the present state of knowledge. The net result, of course, is that we have been provided with a more powerful method of data analysis when contrasted with the initial chi-square test.

Regression Chi-Square Analysis of the Data

Another procedure, a regression chi-square test, has been suggested by Yates (4) and it has been shown to yield results consistent (not identical) with the t-test described above. For this test the null hypothesis is that there is no relation between p_i and X_i. The regression coefficient of p_i on X_i should be a good test criterion . On the null hypothesis each p_i is distributed about the same mean, estimated by $\bar{p}$, with variance $\overline{pq}/n_i$. The regression coefficient b is calculated as

$$b = [\Sigma a_i X_i - (\Sigma a_i)(\Sigma n_i X_i)/N] \ / \ [\Sigma n_i X_i^2 - (\Sigma n_i X_i)^2 / N]$$

Its standard error, S.E. (b), is given by

$$S.E.(b) = [\bar{p} \cdot \bar{q} / \{\Sigma n_i X_i^2 - (\Sigma n_i X_i)^2 / N\}]^{1/2}$$

For the data in Table 3, $b = [168 - (60)(360)/120]/ [1216 - (360)^2/120]$

$$= -0.088235$$

and S.E.(b) = 0.042868. The normal deviate for testing the null hypothesis $\beta = 0$, is $Z = b/S.E.(b) = -2.06$. In this example the regression test gave Z = -2.06 (significant at 0.05 level) while the t-test gave t = -2.08 (118 d.f.). The difference in results arises because the 2 approaches use different large-sample approximations to the exact distribution of Z and t.

When the regression on p_i on X_i is used as a test criterion, it is of interest to examine whether the regression is linear. Armitage (1) has shown that this can be done by first computing chi-square $= \Sigma n_i (p_i - \bar{p})^2 / \bar{p} \cdot \bar{q}$ $= \{\Sigma a_i p_i - A^2/N\} / \bar{p} \cdot \bar{q}$. This chi-square, with (C - 1) d.f., measures the total variation among the C values of p_i. The chi-square for linear regression, with 1 d.f., is found by squaring Z, since the square of a normal deviate has a chi-square distribution with 1 d.f. The difference,

$\chi^2_{(C-1)} - \chi^2_{(1)}$, is a chi-square with (C - 2) d.f. for testing the deviations of the p_i from their linear regression on the X_i. For the school bond example, the total chi-square is 4.72 with 4 d.f., while Z^2 = 4.24 with 1 d.f. Thus the chi-square for the deviations is 0.48 with 3 d.f., which is in agreement with the hypothesis of linearity.

REFERENCES

1. P. Armitage. Biometrics, 11:375(1955).
2. W. G. Cochran. Biometrics, 10:417(1954).
3. G. W. Snedecor and W. G. Cochran. Statistical Methods, 6th ed. The Iowa State University Press, Ames, Iowa (1967).
4. F. Yates. Biometrika, 35:176(1948).

IT IS THE POLICY OF THE SIG=MULTIPLE LINEAR REGRESSION AND OF VIEWPOINTS TO CONSIDER FOR PUBLICATION ARTICLES DEALING WITH THE THEORY AND THE APPLICATION OF MULTIPLE LINEAR REGRESSION. MANUSCRIPTS SHOULD BE SUBMITTED TO THE EDITOR AS AN ORIGINAL, DOUBLE-SPACED TYPED COPY. A COST OF $1 PER PAGE SHOULD BE SENT WITH THE SUBMITTED PAPER. REPRINTS ARE AVAILABLE TO THE AUTHORS FROM THE EDITOR. REPRINTS SHOULD BE ORDERED AT THE TIME THE PAPER IS SUBMITTED AND 20 REPRINTS WILL COST $.50 PER PAGE OF MANUSCRIPT. PRICES MAY BE ADJUSTED AS NECESSARY IN THE FUTURE.

THE UNIVERSITY OF AKRON
AKRON, OHIO 44325

SPANER STEVE
BEHAVIORAL STUDIES
UNIV OF MO-ST LOUIS
ST LOUIS MO 63121

TABLE OF CONTENTS