Multiple Linear Regression Viewpoints
Volume 3, Number 2
October, 1972

A publication of the Special Interest Group on Multiple Linear Regression

of the American Educational Research Association

Table of Contents

SOME FURTHER CONSIDERATIONS OF USING FACTOR REGRESSION ANALYSIS

Isadore Newman

University of Akron

Connett, Houston, and Shaw (1972) presented some advantages and uses of factor regression. I would like to suggest to anyone who is interested in this procedure a program developed by Finn (University of Chicago) who has done some work in this area.

Some Comment on Factor Regression[1]

I. When using factor regression procedures, it is important to keep in mind that if one does not use all of the factors (that is, accounting for 100% of the trace) he may be overlooking a suppressor factor (suppressor variable). Whenever a set of predictor variables is factor analyzed, and accounts for 100% of the trace, and these factors are then used as the set of predictor variables in a multiple regression equation, then the two equations, the one based on the original variables and the other based on the factor variables, the two will make identical predictions (Darlington's Theorem #11 is proof of this statement.) It should be kept in mind, as Connett et al. (1972) pointed out, the factored multiple regression has advantages, such as orthogonality, which allows greater interpretation of the beta weights

[1]Much of these comments are based on a paper by Darlington (1969).

of the multiple regression equations.

II. If one is interested in improving the multiple regression equation by using factor techniques, there is only one way this can be done. That is if the number of factors that are used is less than the number of original variables. This will increase the df and also possibly decrease shrinkage-estimates. Because of this, some researchers have used only the few factors that account for the "greatest" amount of the factored trace. However, when this is done one may be losing information that can account for criterion variance by eliminating a factor that accounts for very little trace of the factored matrix but is highly correlated with the criterion scores.

III. An additional advantage of using factor scores is that when a matrix is factored much of the error variance tends to be distributed in the factors that account for the least variance. Therefore, one of the possible by-products of using factor scores which account for most trace variance as predictors is the likelihood of increasing reliability (therefore decreasing shrinkage).

Using only the factors that account for most of the trace should be avoided when the predictor variables that are being factored are likely to be highly reliable. Some examples of such variables are: height, weight, religion, sex, income, age, etc. Under these conditions a variable that accounts for little of the trace variance may be a good and highly reliable predictor of criterion variance.

IV. Connett et al. (1972) suggested that using factor regression should be of help when developing multiple regression equations for exploratory purposes. If one combines this factor regression procedure with stepwise procedures, some of the problems discussed in Sections I-III, can be considerably reduced. When stepwise procedures are employed, one should always use all the factors needed to account for 100% of the trace of the predictor variables.

As most of you are aware, the two basic stepwise procedures are stepup and stepdown. The advantages of each will not be discussed here but should be considered when one is choosing a stepwise procedure. One concept of "tolerance" is not as widely familiar. "Tolerance" is a measure of how much a particular variable in an equation is a linear combination of the other variables in that equation. When there exists a perfect linear relationship, the "tolerance" will equal zero. If the "tolerance" of a particular variable is one, this means that the variable is adding an additional new dimension to the predictor equation which the other variables are not accounting for. The amount of additional variance accounted for by adding any variable is the product of the normalized regression coefficient squared and the "tolerance" of that variable. This "tolerance" procedure is used in the Stepwise Regression Program, in the Statistical Package for the Social Sciences (SPSS). When the factor regression procedure is used each factor will have, by definition, a "tolerance" equal to one. It can be seen how this "tolerance" procedure may allow one to use the original set of predictor variables and get much of the same information one would obtain by using factor regression procedure.

REFERENCES

Connett, Houston, and Shaw. "The Use of Factor Regression in Data Analysis", Multiple Linear Regression Viewpoints, 1972 (March) vol. 2, No. 4, 46-49.

Darlington, R.B. "Multiple regression in psychological research and practice", Psychological Bulletin, 1969, Vol. 69, No. 3, 161-182.

PROOF THAT THE DEGREES OF FREEDOM FOR THE TRADITIONAL METHOD OF CALCULATING ANALYSIS OF COVARIANCE AND THE MULTIPLE REGRESSION METHOD ARE EXACTLY THE SAME

Isadore Newman
Jon Fry
University of Akron
Akron, Ohio

Recently, a variety of publications (O'Conner, 1972; Kelly, Beggs, McNeil, 1969; Williams, Maresh, Peebles, 1972) have indicated that multiple regression is a very efficient method of calculating analysis of covariance. However, it is still unclear if the analysis of covariance, using the multiple regression approach suggested by Kelly, et al (1969), produces exactly the same result as the more traditional method of calculating analysis of covariance. The traditional method uses the residuals left after the variance of the criterion that can be accounted for by the covariate is subtracted. When comparing these two approaches, one may find differences due to incorrect calculations of the degrees of freedom (df) since it is easy to make an error in calculating the df when comparing the above two methods.

The following four models are presented to prove that the df's for both approaches are exactly the same:

Where: Y_1 = Posttest scores
X_1 = 1 if the S is in Group 1: 0 otherwise

X_2 = 1 if the S is in the Control Group; 0 otherwise
X_3 = Pretest scores
$E_1 = Y_1 - \hat{Y}_1$ (for Model 1)
E_2 = Residuals left after the variance that the pretest score accounts for is subtracted from the posttest score ($Y_1 - \hat{Y}_1$ for Model 2)

U = Unit vector; 1 if the S is in the sample; 0 otherwise
a_0, a_1, a_2, a_3, = Partial regression weights

Let us also assume that N (the number of Ss) = 100.

Model 1: $Y_1 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + E_1$
Restriction $a_1 = a_2$
Model 2: $Y_1 = a_0U + a_3X_3 + E_2$
Model 3: $E_2 = a_0U + a_1X_1 + a_2X_2 + E_3$
Restriction $a_1 = a_2$
Model 4: $E_2 = a_0U + E_4$

If one calculated the df needed for the F-ratio for testing Model 1 against Model 2, the correct df would be:

df 1 = 3-2 = 1
df 2 = 100-3 = 97

where: df 1 = the number of linearly independent vectors in the full model (Model 1) minus the number of linearly independent vectors in the restricted model (Model 2).

df 2 = the number of independent observations minus the number of linearly independent vectors in the full model.

If the df for testing Model 3 against Model 4 is calculated in exactly the same way, one can easily make an error as follows:

df 1 = 1
df 2 = 100-2 = 98

This is an error. The following is a proof that the second set of df is incorrect since N is defined as the number of independent observations.

Theorum: In any sample of N observations, if multiple linear regression is performed on the model $Y_i = \sum_{j=1}^{m} a_j X_{ij} + E_i$ given the least squares coefficients a_j, $j = 1, 2, \ldots n$, then, given all values of the independent variables X_{ij}, and the coefficients a_j, and any N-m of the residuals E_i, the remaining m residuals can be computed.

Proof:

(1) $E_i (Y_i - a_j X_{ij}) = 0$ for any j

(2) $Y_i = e_i + E_j a_j X_{ij}$ (model equation)

substituting (2) into (1)

(3) $E_i (e_i + E_j a_j X_{ij} - a_j X_{ij}) = 0$ for any j each value of j gives a linear equation in the missing e_i. Since $j = 1, 2, \ldots n$ and there are m missing e_i, (3) gives m linear equations in m unknowns. Hence, the number of independent observations in the vector of residuals is N-m, not N.

The major point of this paper is that the df for both methods of calculating analysis of covariance are exactly the same. This does not mean that we believe the results will be exactly the same, for that is an emperical question (as Keith McNeil would say). To determine that, a Monte Carlo study is being planned. If you have any suggestions as to how we may best proceed, we would appreciate your comments.

REFERENCES

Kelly, F. J., Beggs, D. L., McNeil, K. A., Eichelberger, T., and Lyon, J. Research Design in the Behavioral Sciences, Multiple Regression Approach. Southern Illinois University Press, Carbondale, 1969.

O'Conner, E. J. "Extending Classical Test Theory to the Measurement of Change," Review of Educational Research, Vol. 42, No. 1, Winter, 1972.

Williams, J. D., Maresh, R. T., and Peebles, J. D. "A Comparison of Raw Gain Scores, Residual Gain Scores, and the Analysis of Covariance with Two Modes of Teaching Reading," Multiple Linear Regression Viewpoints, Vol. 3, No. 1, May, 1972.

CONDITIONS FOR NO SECOND ORDER INTERACTION IN MULTIPLE LINEAR REGRESSION MODELS FOR THREE FACTOR ANALYSIS OF VARIANCE

By M. A. BREBNER

University of Calgary, Alberta

PART I - PICTORIAL OR GEOMETRIC APPROACH

Very little discussion is given on the subject of second order interaction in analysis of variance. Many texts, if they even discuss the matter, state the general mathematical condition (Kendall and Stuart (1) page 36)

$$\sum_{j=1}^{r} \sum_{k=1}^{s} \sum_{l=1}^{t} (\theta_{jkl})^2 = 0 \qquad \ldots 1(a)$$

where

$$\theta_{jkl} = \mu_{jkl} - (\mu_{jk*} + \mu_{j*l} + \mu_{*kl}) + (\mu_{j**} + \mu_{*k*} + \mu_{**l}) - \mu_{***} \qquad \ldots 1(b),$$

and dismiss the details as cumbersome or complex. Meyers (2) gives a set of conditions on the cell means, but does not explain that only a selected subset of the conditions are required and that the other conditions are redundant (duplicate the subset of conditions). Winer (3) endeavours to give a pictorial explanation of no second order interaction which on closer examination is found to deal with only one very special case. This will be discussed in section 1 where a three factor 2*2*2 design is considered. In section 2 the results of section 1 will be generalized to cases where more than two levels occur in a factor.

To avoid complicating the description, the following assumption will be added to the normal statistical requirements:

1) The number of scores per cell is the same for each cell, namely, n.

2) The scores are not subject to numerical error.

The difficulties and complications that can arise in two factor designs if assumption 1) is violated have been discussed by White (4). A good text on industrial statistics should discuss the effect of assumption 2).

1. In this section we will discuss the condition for no second order interaction in a three factor 2*2*2 design. Firstly, we will formally state the two multiple linear regression models required, and will subsequently give a primarily pictorial (geometric) explanation of the restriction required for model 2. The three factors will be labelled A, B and C, and individual cells will be referred to by factors and levels. For example, A1B2C2 will refer to the cell at level 1 of factor A, level 2 of factor B and level 2 of factor C.

Model 1 No restriction

$$\underline{y} = a_1\underline{x}_1 + a_2\underline{x}_2 + a_3\underline{x}_3 + a_4\underline{x}_4 + a_5\underline{x}_5 + a_6\underline{x}_6 + a_7\underline{x}_7 + a_8\underline{x}_8 + \underline{e}_1 \qquad \ldots 2(a)$$

Model 2 Restriction for no second order interaction

$$(a_4 - a_3) - (a_2 - a_1) = (a_8 - a_7) - (a_6 - a_5) = I \qquad \ldots 2(b)$$

$$\underline{y} = a_1\underline{x}_1 + a_2\underline{x}_2 + a_3\underline{x}_3 + (I + a_3 + (a_2 - a_1))\,\underline{x}_4$$
$$+ a_5\underline{x}_5 + a_6\underline{x}_6 + a_7\underline{x}_7 + (I + a_7 + (a_6 - a_5))\underline{x}_8 + \underline{e}_I$$

$$\text{or } \underline{y} = a_1(\underline{x}_1 - \underline{x}_4) + a_2(\underline{x}_2 + \underline{x}_4) + a_3(\underline{x}_3 + \underline{x}_4) + I(\underline{x}_4 + \underline{x}_8)$$
$$a_5(\underline{x}_5 - \underline{x}_8) + a_6(\underline{x}_6 + \underline{x}_8) + a_7(\underline{x}_5 + \underline{x}_8) + \underline{e}_I \qquad \ldots 2(c)$$

where the subscript 1 refers to the cell A1B1C1,

"	"	2	"	"	"	"	A2B1C1,
"	"	3	"	"	"	"	A1B2C1,
"	"	4	"	"	"	"	A2B2C1,
"	"	5	"	"	"	"	A1B1C2,
"	"	6	"	"	"	"	A2B1C2,
"	"	7	"	"	"	"	A1B2C2,
"	"	8	"	"	"	"	A2B2C2.

The column vectors $\underline{y}$ and $\underline{x}_i$, $i = 1,2,\ldots,7,8$ are of length $2_*2_*2_*n$

Scores from all cells are arranged systematically in the vector $\underline{y}$.

The vectors $\underline{x}_i$ are defined as follows:

$$\text{an element of the vector } \underline{x}_i = \begin{cases} 1 & \text{if the corresponding element of } \underline{y} \text{ is a score from the cell } i. \\ 0 & \text{otherwise.} \end{cases}$$

The parameters a_i and I are determined by a least squares fit which minimizes $\underline{e}_1^T\underline{e}_1$ in model 1 and $\underline{e}_I^T\underline{e}_I$ in model 2, where $\underline{e}^T$ denotes the transpose of the column vector $\underline{e}$ (that is a row vector with the same elements as $\underline{e}$) . For the reader not familiar with matrix notation the quantity $\underline{e}_1^T\underline{e}_1$ is just the error sum of squares for model 1. Note the value of a_i computed by model 1 will be the mean of the scores from cell i .

To test for a significant second order interaction the following F statistic with $(r-1)(s-1)(t-1)$ and $(rst(n-1))$ degrees of freedom is computed

$$F = \frac{(ESS_I - ESS_1)/(r-1)(s-1)(t-1)}{(ESS_1)/(rst(n-1))} \qquad \ldots 2(d)$$

where $ESS_1 = \underline{e}_1^T \underline{e}_1$,

$ESS_I = \underline{e}_I^T \underline{e}_I$,

$r, s,$ and t are the number of levels for factors A, B and C respectively (in this case $r=s=t=2$) , and n is the number of observations per cell.

To simplify the later discussion a measure of first order interaction for a 2*2 design will be given. For this purpose consider the C1 level in the 2*2*2 design under discussion. This involves the cells with

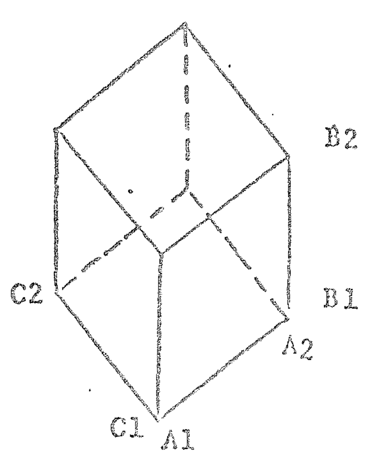

Fig. 1(a)

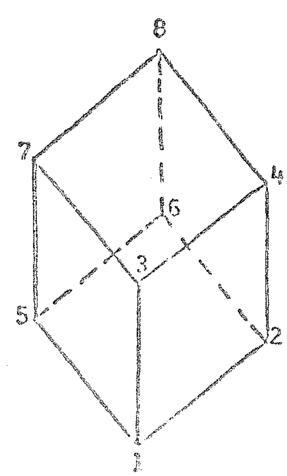

Fig. 1(b)

index i = 1,2,3,4 . A measure I_{1234} of the first order interaction over the face 1234 of the cube given in figures 1(a) and 1(b) is defined by the equation

$$I_{1234} = (a_4 - a_3) - (a_2 - a_1) \ .$$

It should be noted at this point that the orientation of the corner points of the face must be consistently ordered when evaluating the right hand side of the equation. For example:

$$(a_3 - a_4) - (a_1 - a_2) \quad = \quad - I_{1234}$$

does not equal I_{1234}. Similarly

$$(a_2 - a_1) - (a_4 - a_3) \quad = \quad I_{3412} \neq I_{1234}$$

We can now state the restriction or condition for no second order interaction in a 2*2*2 design as follows:

$$I_{1234} \quad = \quad I_{5678} \quad = \quad I \qquad \ldots\ 3(a)$$

$$\text{or} \quad (a_4 - a_3) - (a_2 - a_1) \quad = \quad (a_8 - a_7) \quad - \quad (a_6 - a_5) \quad = \quad I \qquad \ldots\ 3(b)$$

This restriction states that the first order interaction over the face 1234 equals the first order interaction over the perpendicularly opposite face 5678 or the first order interaction is constant over the perpendicularly opposite C1 and C2 faces.

The faces in a 2*2*2 design are just "squares", but if more levels in a factor are considered a face consists of several "squares". (see figures 4(a) and 4(b)). For convenience in section 2 we will refer to these "squares" as panels. For example, in 4*3*3 design the C1 face would contain (4-1)*(3-1) = 6 panels (see figure 5).

Returning to our simple 2*2*2 design, we could obtain two further restrictions by equating first order interactions over

(i) the A1 and A2 faces or the panels 1537 and 2648

and (ii) the B1 and B2 faces or the panels 1256 and 3478

This gives the equations

$$(a_7 - a_3) - (a_5 - a_1) = (a_8 - a_4) - (a_6 - a_2) = I_A \quad \ldots \quad 4(a)$$

and

$$(a_6 - a_5) - (a_2 - a_1) = (a_8 - a_7) = (a_4 - a_3) = I_B \quad \ldots \quad 4(b)$$

For uniformity of notation we could replace I in equation 3(b) by I_C as the equation relates to interactions over the C1 and C2 faces.

$$(a_4 - a_3) - (a_2 - a_1) = (a_8 - a_7) - (a_6 - a_5) = I_C \quad \ldots \quad 4(c)$$

Though I_A, I_B and I_C are not equal the relations between the parameters a_i in 4(a), (b) and (c) are equivalent. For example, consider 4(a)

$$(a_7 - a_3) - (a_5 - a_1) = (a_8 - a_4) - (a_6 - a_2)$$

Add $(a_4 - a_2 - a_7 + a_5)$ to each side. The terms in a_7 and a_5 cancel on the left and the terms in a_4 and a_2 cancel on the right of the equation giving the equation

$$-a_3 - (-a_1) + a_4 - a_2 = a_8 - (a_6) - a_7 + a_5$$

or

$$(a_4 - a_3) - (a_2 - a_1) = (a_8 - a_7) - (a_6 - a_5)$$

which is exactly the relation between the parameters a_i given by equation 4(c) It can be similarly shown that relation 4(b) is equivalent to the relations 4(a) and 4(c). Hence it is only necessary, and in fact sufficient, to use one of the three relations obtained by considering the faces of the cube given in figures 1(a) and (b). That the relations given above do guarantee no second order interaction will be proved mathematically in part II of this paper.

We now consider the pictorial explanation given in Winer (3) on pages 182 to 184.

Winer (3) gives an example like example 1 below.

C 1

	B 1	B 2
A 1	2	8
A 2	2	4

C 2

	B 1	B 2
A 1	4	10
A 2	4	6

Table 1 - Cell Means for example 1

The means for this example are shown diagramatically in figures 2(a) and (b).

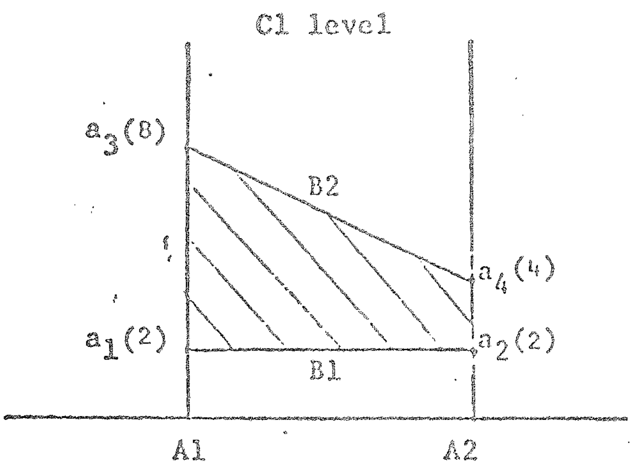

Fig. 2(a)

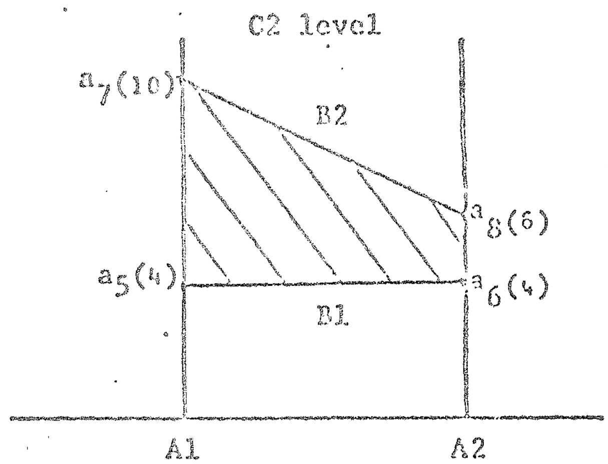

Fig. 2(b)

Note the shaded areas in figures 2(a) and 2(b) are similar. This similarity of shape, Winer gives as a condition for no second order interaction. Certainly the cell means for example 1 satisfy the condition 4(c)

$$(4 - 8) - (2 - 2) = (6 - 10) - (4 - 4)$$

Now consider the following set of data.

C 1

	B 1	B 2
A 1	6	4
A 2	2	5

C 2

	B 1	B 2
A 1	4	6
A 2	2	9

Table 2 - Cell Means for example 2

We illustrate the means for example 2 diagramatically in figures 3(a) and (b).

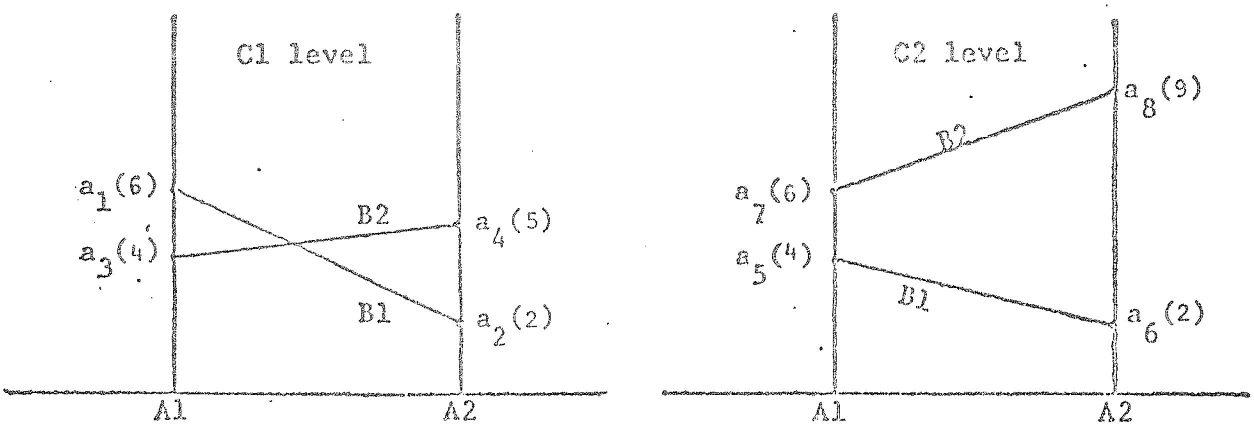

Fig. 3(a) Fig. 3(b)

No similar shapes can be observed in figures 3(a) and (b), but the data does satisfy the restriction 4(c).

$$(5 - 4) - (2 - 6) = (9 - 6) - (2 - 4)$$

There is no second order interaction for the data of example 2, even though "Winer's condition" is violated. The "Winer condition" is a special case of the completely general condition given in this paper.

2. How do these ideas generalize in the case of more than two levels for each factor? There are now many "cubes" of the type shown in

figures 1(a) and (b). This is illustrated by figures 4(a) and (b), and figure 5. For simplicity we now consider a 3*2*2 design shown diagramatically in figures 4(a) and (b).

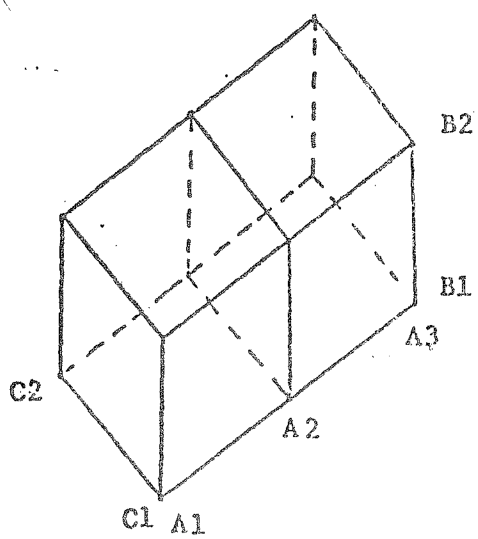

Fig. 4(a)

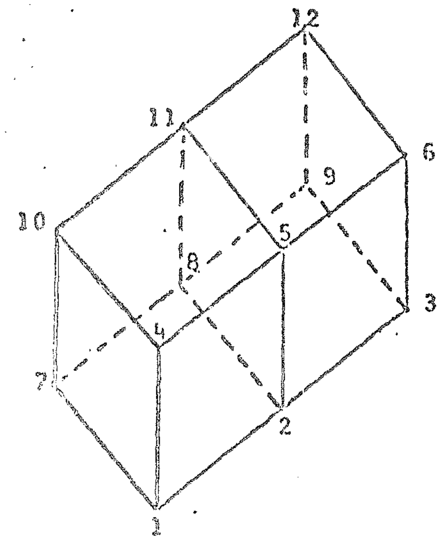

Fig. 4(b)

The rectangular block in figures 4(a) and (b) is split into two 2*2*2 cubes by level 2 of A. It was noted in section 1 that a cube produces 3 restrictions, 2 of which are redundant. The two cubes will produce 6 restrictions, 4 of which are redundant. There are three alternative ways of selecting 2 non-redundant restrictions as follows:

(i) Equate the BC first order interactions on the faces $A1$, $A2$ and $A3$. That is

$$I_{1,7,4,10} = I_{2,8,5,11} = I_{3,9,6,12} = I_A \qquad \ldots 5(a)$$

(ii) Equate the AC first order interactions for the two panels on the $B1$ and $B2$ faces. That is

$$I_{1,2,7;8} = I_{4,5,10,11} = I_{B_1} , \qquad \ldots 5(b)$$

and

$$I_{2,3,8,9} = I_{5,6,11,12} = I_{B_2} . \qquad \ldots 5(c)$$

(iii) Equate the *AB* first order interactions for the two panels on the *C*1 and *C*2 faces. That is

$$I_{1,2,4,5} = I_{7,8,10,11} = I_{C_1}\,, \qquad \ldots 5(d)$$

and

$$I_{2,3,5,6} = I_{8,9,11,12} = I_{C_2}\,. \qquad \ldots 5(e)$$

Any one of the pairs (i), (ii) or (iii) are necessary, and in fact sufficient, conditions for no second order interaction in a 3*2*2 design. It can be noted here that "Winer's condition" is sufficient, but not necessary.

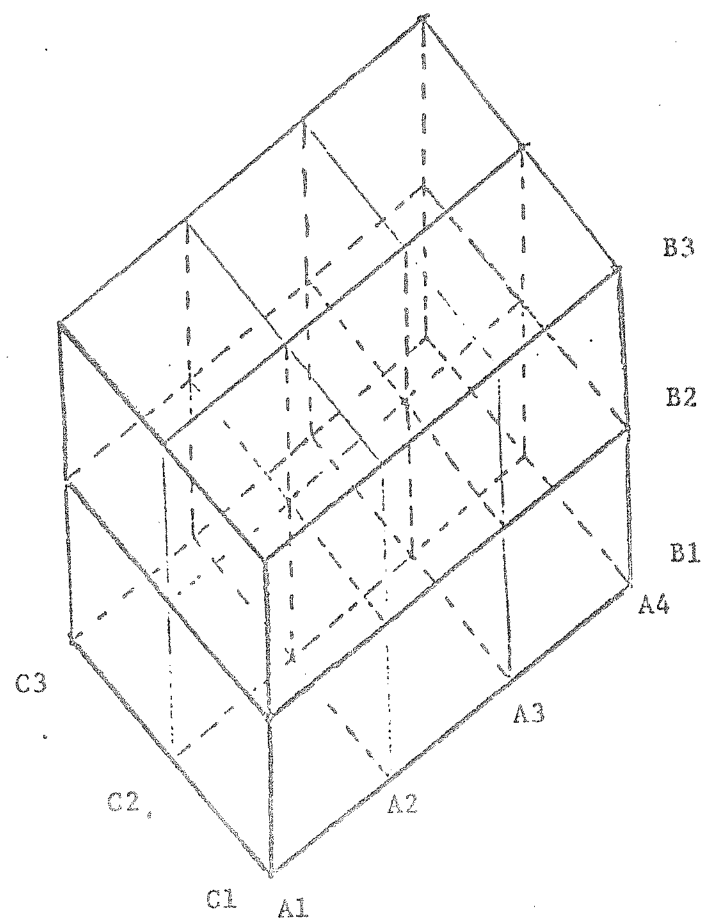

Fig. 5

A general rule emerges. Take all panels on the face A1. Then equate the first order interaction on each of these panels with the interaction on the corresponding panels at every other level of A. If we are considering a r*s*t design there are (s-1)(t-1) panels on the A1 face. Each of these panels will produce (r-1) restrictions. Hence there is a total of (r-1)(s-1)(t-1) restrictions. For the case of 4*3*3 design there are (4-1)(3-1)(3-1)=16 restrictions. We could equally well have taken the (r-1)(t-1) panels on the B1 face and equate interactions on the corresponding panels at the (s-1) other levels of factor B. Again (r-1)(s-1)(t-1) restrictions are generated. This is exactly the number of cubes in the design. For each cube there is one restriction from the set generated by A1 and one restriction from the set of restrictions generated by B1. But the restriction given by two A faces of a cube and the restriction given by two B faces are equivalent. Hence the first set of restrictions is duplicated by the second set of restrictions. A third set of restrictions based on C1 and other levels of C will again duplicate the first set of restrictions and consequently duplicate the second set of restrictions. Let us finally give the models for a 3*2*2 design.

Model 1 No restriction

$$\underline{y} = a_1\underline{x}_1 + a_2\underline{x}_2 + a_3\underline{x}_3 + a_4\underline{x}_4 + a_5\underline{x}_5 + a_6\underline{x}_6 + a_7\underline{x}_7 + a_8\underline{x}_8 + a_9\underline{x}_9 + a_{10}\underline{x}_{10} + a_{11}\underline{x}_{11} + a_{12}\underline{x}_{12} + \underline{e}_1 \qquad \ldots 6(a)$$

Model 2 Restriction for no second order intersection

$$(a_{10} - a_4) - (a_7 - a_1) = (a_{11} - a_5) - (a_8 - a_2) = (a_{12} - a_6) - (a_9 - a_3) = I_A \qquad \ldots 6(b)$$

$$\underline{y} = \sum_{i=1}^{9} a_i\underline{x}_i + \{I_A + a_4 + (a_7 - a_1)\}\,\underline{x}_{10} + \{I_A + a_5 + (a_8 - a_2)\}\,\underline{x}_{11} + \{I_A + a_6 + (a_9 - a_3)\}\,\underline{x}_{12} + \underline{e}_I$$

$$
\begin{aligned}
\text{or } \underline{y} = {} & a_1(\underline{x}_1 - \underline{x}_{10}) + a_2(\underline{x}_2 - \underline{x}_{11}) + a_3(\underline{x}_3 - \underline{x}_{12}) \\
& + a_4(\underline{x}_4 + \underline{x}_{10}) + a_5(\underline{x}_5 + \underline{x}_{11}) + a_6(\underline{x}_6 + \underline{x}_{12}) \\
& + a_7(\underline{x}_7 + \underline{x}_{10}) + a_8(\underline{x}_8 + \underline{x}_{11}) + a_9(\underline{x}_9 + \underline{x}_{12}) \\
& + I_A(\underline{x}_{10} + \underline{x}_{11} + \underline{x}_{12}) + \underline{e}_I \qquad \ldots 6(c)
\end{aligned}
$$

where the indexing is consistent with figures 4(a) and (b), and the definition of the vectors $\underline{x}_i$ and $\underline{y}$ is essentially the same as for equations 2(a) and (c). Note the parameters a_i computed by least squares for Model 1, will not in general be equal to the parameters a_i computed for Model 2.

A final word of warning is that second and higher order interactions must be interpreted with great care, if meaningless or erroneous conclusions are not to be drawn from research data.

REFERENCES

1) Kendall, M.G., and Stuart, A., The Advanced Theory of Statistics Vol. III 3rd edition, Charles Griffin and Co. 1969.

2) Myers, J.L., Fundamentals of Experimental Design, Allan and Bacon 1966.

3) Winer, B.J., Statistical Principles in Experimental Design, McGraw-Hill 1962.

4) White, R., M. Ed. Thesis May 1971, University of Calgary

A Note on the Independent Variance of Each Criterion in a Set

Isadore Newman, University of Akron
and
Keith McNeil, Southern Illinois University at Carbondale

Frequently a researcher is interested in a number of criterion variables which may not be uncorrelated with each other. An example of such criterion variables may be History GPA Y_1, Math GPA Y_2, and Reading GPA Y_3. The chances are that these variables are likely to be significantly correlated with each other.

If one developed a regression equation trying to predict Y_1, (History GPA) from socioeconomic status (X_1) it would be presented in the form of Model 1.

Model 1 $Y_1 = a_0U + a_1X_1 + E_1$

One can test to see how much of the criterion variance is accounted for by looking at the R^2. However, if the researcher is interested in determining how much of the criterion variance X_1 can account for of Y_1, that is independent of Y_2 and Y_3, he can test Model 2 against Model 3.

Model 2 $Y_1 = a_0U + a_1X_1 + a_2Y_2 + a_3Y_3 + E_2$

Model 3 $Y_1 = a_0U + a_1Y_2 + a_2Y_3 + E_3$

By testing Model 2 against Model 3 we are testing the amount of variance of Y_1 that is independent of Y_2 and Y_3, that can be accounted for by X_1.

If we follow this exact procedure for predicting Y_2 and Y_3, the three F tests that would be run, for Y_1, Y_2 and Y_3, would be independent of each other. Therefore the probability of making a Type I error for each test would be equal to α (assuming the underlying assumptions of normality and homogeneity of variance are met).

Another way of looking at this is if one did not covary the additional criterion variables, one would be required to use some multiple comparison correction since the three F ratios for the three criterion variables would not be independent of each other. For example, if we use the concept $\frac{\alpha}{N}$[1] (where N = number of nonindependent F-tests calculated and α = the desired level we want to hold constant) for this purpose then the N, for the first case, where the criterion variables were covaried would be equal to one since all three F ratios would be independent of each other. However, in the second case where the criterion variables are not covaried, the F ratios would not be independent and, therefore, N would be equal to three.

If there is more than one predictor variable, X_1(Socioeconomic status), X_2(Sex), X_3(Years in school) and X_4(Age), and we were interested in determining how much of the independent variance of the criterion each of these variables accounted for, one would test Model 4 against Model 5, Model 4 against 6, Model 4 against 7, and Model 4 against 8. A similar procedure would be followed for criterion variables 2 and 3. The N then needed for $\frac{\alpha}{N}$ multiple comparison correction would then be 12 (four for testing Y_1, four for Y_2 and four for testing Y_3 since Y_1, Y_2 and Y_3 are not independent of each other.

Model 4 $\quad Y_1 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + E_4$

Model 5 $\quad Y_1 = a_0U + a_2X_2 + a_3X_3 + a_4X_4 + E_5$

Model 6 $\quad Y_1 = a_0U + a_1X_1 + a_3X_3 + a_4X_4 + E_6$

Model 7 $\quad Y_1 = a_0U + a_1X_1 + a_2X_2 + a_4X_4 + E_7$

Model 8 $\quad Y_1 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + E_8$

---

[1]This formula and proof has been presented in Multiple Linear Regression Viewpoints Vol. 2, No. 3, Jan. 1972, by Newman & Fry, p. 36-39.

However, if models similar to Models 9 and 10 which covaried the additional criterion variables were used following the same procedure that was in the above test for each of the criterion variables: Y_1, Y_2 and Y_3, the N needed for the $\frac{\alpha}{N}$ multiple correction, would be 4 instead of 12.

Model 9 $\quad Y_1 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5Y_2 + a_6Y_3 + E_9$

Model 10 $\quad Y_1 = a_0U + a_2X_2 + a_3X_3 + a_4X_4 + a_5Y_2 + a_6Y_3 + E_{10}$

As suggested above, by covaring other nonindependent criteria that we are interested in, we can actually treat each test of each criterion as if the criterion variables were independent of each other. Of course, testing Model 9 versus 10 answers a different question than does testing Model 4 versus 5. The question one wants to ask should always dictate the models used. What we are saying here is that if one is interested in accounting for the independent piece of a criterion variable's variance, then the procedures outlined here should be used.

MEMBERSHIP LIST

William Adrian
Special Assistant to the Chancellor
University of Denver
Denver, CO 80210

Earl A. Allusi
Director
Performance Research Laboratory
University of Louisville
Louisville, KY 40208

Arnold G. Ashburn
Deputy Assistant Superintendent
Research & Evaluation
Dallas Independent School District
3700 Ross Avenue
Dallas, TX 75204

Richard Arakaki
Hawaii State Department of Education
Honolulu, Hawaii

Sylvia Auton
%Dept. of Educational Measurement
and Statistics
University of Maryland
College Park, MD

Paul Barbuto, Jr.
Box 37
Teachers College
Columbia University
New York, NY 10027

Richard Beeson
Research Methodology
St. Louis University
211 N. Grand
St. Louis, MO 63103

Donald L. Beggs
Dept. of Guidance & Ed. Psych.
Southern Illinois University
Carbondale, IL 62901

Katherine Bemis
S. W. Cooperative Education Laboratory
Albuquerque, New Mexico

William Beusse
AFHRL
Wright Patterson AFB
Maryland

E. H. Biekking
Box 12524
University Station
Gainesville, Florida 32601

James Bolding
Box 339
Fayetteville, Arkansas 72701

Robert Bottenberg
Air Force Personnel Research Division
AFHRL, Box 1557
Lackland AFB, Texas 78236

Michael Brebnger
Dept. of Ed. Psych.
University of Calgary
Calgary 44, Alberta
CANADA

William Brockhaus
University of Southern California
Graduate School of Business Administration
Management Department
Hoffman Hall Office 701N
Los Angeles, CA 90007

William Brookshire
North Texas State University
P. O. Box 13841
Denton, TX 76203

K. Brown
Apt. 159 Villa de Palmer
5811 Atlantic Blvd.
Jacksonville, Florida 32202

Robert Brownlee
CTB/McGraw-Hill
Del Monte Research Park
Monterey, CA 93940

David Buckholdt
10646 St. Charles Rock Rd.
St. Louis, MO 63074

Jeff Bulcock
Department of Educational Foundations
Memorial University of Newfoundland
St. John's Newfoundland
CANADA

Leigh Burstein
744 Coleman Avenue #E
Menlo Park, CA 94025

Leonard S. Cahen (Dr.)
Educational Testing Service
Division of Psychological Studies
Princeton, New Jersey 08540

Dale Carlson
Office of Program Evaluation
Department of Education
State Education Building
721 Capitol Mall
Sacramento, CA 95814

Donald Cegala
241 Dalton Court
Tallahassee, Florida 32304

Robert Coldiron
Educational Research Associate
Dept. of Public Instruction
Harrisburg, PA 17126

Gary Coles
Associate Research Scientist
American Institutes for Research
Center for Research and Evaluation in
the Applications of Technology in Education
P.O. Box 1113
Palo Alto, CA 94302

Arnold Coltvet
Iowa Central Community College
330 Avenue M
Fort Dodge, IA 50501

William Connett
Dept. of Research & Statistical Methodology
University of Northern Colorado
Greeley, CO 80631

John Convey
Florida State University
Tallahassee, Florida

Louise Corman
Research Institute for Educational
Problems
12 Maple Avenue
Cambridge, Massachusetts 02139

Hal Corson
Research Associate
Miami Dade Junior College
11011 S. W. 104 Street
Miami, Florida 33156

Paul T. Costa
Harvard University
Dept. of Social Relations
William James Hall
Cambridge, Massachusetts 02138

Laura R. Crane
Chicago Board of Education
Dept. Govt. Funded Room 1130
228 N. LaSalle
Chicago, IL

Carl Crosswhite
University of Northern Colorado
Greeley, CO 80631

William Denton
7777 Forney Road
Dallas, TX 75227

William Donaldson
Delaware Rd., Piney Ridge
Pine Grove Mills, PA 16868

Vern Dravland
Professor of Education
Coordinator of Educational Research
University of Lethbridge
Lethbridge, Alberta
CANADA

Charles Eberly
Office of Evaluative Services
239 S. Kedzie Hall
Michigan State University
East Lansing, MI 48823

Tony Eichelberger
Learning Research & Development Center
University of Pittsburgh
Pittsburgh, PA

Patricia Elmore
Counseling & Testing Center
Southern Illinois University
Carbondale, IL 62901

Claire Ernhart
10670 St. Charles Road Rd.
St. Ann, MO 63074

Avigdor Farine
Faculte des Sciences de l'Education
Section d'Administration Scolaire
University de Montreal
Case postal 6128
Montreal 3, Quebec
CANADA

Garrett R. Foster
Assistant Professor
405 Education
Florida State University
Tallahassee, Florida 32306

Paul Games
420 Social Sciences
University Park, PA 16802

Robert Gardner
1305 6th Street
Greeley, CO 80631

Sara Joy Graham
307-314 Southwest 16th Avenue
Gainesville, Florida 32601

Gordon Graves
Fresno Unified School District
2348 Maripisa
Fresno, CA 93721

Vincent Greavey
Boston College
Boston, MA

Beatrice Green
East Street
Stockbridge, MA

Barry Greenberg
8604 S. W. 102 Place
Miami, Florida 33143

Hugh Greenup
5406 Rhea Avenue
Tarzana, CA 91356

Willa Gupta
Data Analyst, UCLA
Educational Preschool Language
1868 Greenfield Avenue
Los Angeles, CA 90025

Donald Haefele
210-C Arps Hall
Tho Ohio State University
Columbus, Ohio 43210

Ofelia Halasa
Div. of Research and Development
Cleveland Public Schools
Cleveland, Ohio

Robert Hale
Research Methodology
221 N. Grand
St. Louis University
St. Louis, MO 63130

Mrs. Irma Halfter
Vice-Pres., Analytic Studies
DePaul University
25 East Jackson Blvd.
Chicago, IL 60604

Ronald S. Halinski
Assistant Prof. of Education
Department of Education
Illinois State University
Normal, IL 61761

Marvin Halldorson
School of Business
University of Northern Colorado
Greeley, CO 80631

Joe B. Hansen
Education Service Center
Region 13
Austin, TX

Beatrice Harris
Yeshiva University
55 Fifth Avenue
New York, NY 10003

Joseph Harrison
Math Coordinator, Exp. in Higher Ed.
13200 St. Louis Avenue
East St. Louis, IL 62201

Jack Haynes
Dept. of Psychology
Box 13587, N. T. Station
Denton, TX 76203

Mrs. Letitia Heil
204 Briarton Lane
Crystal City, MO 63019

Beatrice Heimerl
Dept. of Research & Statistical Methodology
University of Northern Colorado
Greeley, CO 80631

Dennis Hein
Augustana College
Sioux Falls, SD 57102

Carl Helwig
School of Education
Old Dominion University
Norfolk, VA 23505

John Hemmeter
Indiana University
Bloomington, Indiana

Veron Hendrix
Professor
221 Borton Hall
University of Minnesota
Minneapolis, MN 55455

James D. Hennes
Program Evaluation Center
Univ. of Missouri Medical Center
201 Lewis Hall
Columbia, MO 65201

Sarah Hervey
College of Education
University of Arizona
Tucson, Arizona 85721

Thomas L. Hick
Director, Child Study Center
Campus School
State University College
New Paltz, NY 12561

Vynce Hines
1220 S. W. Ninth Road
Gainesville, Florida 32601

Jack I. Hoffman
4845 Jerome Avenue
Skokie, IL 60076

James Hogge
School of Education
Box 512
George Peabody
Nashville, TN 37200

Samuel R. Houston
Assistant Professor
Dept. of Research & Statistical Methodology
University of Northern Colorado
Greeley, CO 80631

Carl Huberty
325 Aderhold
University of Georgia
Athens, GA 30601

Brad Huitman
Psychology Department
Western Michigan University
Kalamazoo, MI 49001

Janet Carol Hyde
1048 M Graduate House
West Lafayette, Indiana 47906

Earl Jennings
School of Education
University of Texas
Sutton Hall 6
Austin, TX 78721

Lary Johnson
Research and Evaluation
807 N. E. Boradway
Minneapolis, MN 55413

Paul Jones
Research and Development Division
American College Testing Program
P.O. Box 168
Iowa City, IA 52240

Thomas Jordan
EDAP Program Director
CEMREL
10646 St. Charles Rock Road
St. Ann, MO 63074

Daniel C. Kau
Dept. of Ed. Psych.
University of Illinois
Urbana, IL 61801

Francis Kelly
Professor of Educational Psychology
Educational Research Bureau
Southern Illinois University
Carbondale, IL 62901

John Kennedy
229 Ramseyer Hall
The Ohio State University
Columbus, Ohio 43210

Raymond Kieft
Director
Institutional Research
Central Michigan University
Mt. Pleasant, MI 48858

F. J. King
403 Education Building
Institute of Human Learning
Florida State University
Tallahassee, Florida 32306

Alan Klass
Milwaukee Luthern High School
9700 West Grantosa Drive
Milwaukee, WI 53222

J. A. Klock
2226 Mercer Circle So.
Jacksonville, Florida 32317

Harold B. Knight
Director, Education Research
Box 68, Southern Station
University of Southern Mississippi
Hattiesburg, Mississippi 39401

Patricia Knox
6350 N. Lake Drive
Milwaukee, WI 53217

Ronald Koback
1410 South Dixie Highway
Coral Gables, Florida 33146

Richard Kohr
500 Winand Drive
Harrisburg, PA 17109

Janos Koplyay
Personnel Research Division
Air Force Human Resource Laboratory (HRPS)
Lackland AFB, Texas 78236

Conrad Krauft
2371 Ora Drive
Fayetteville, Arkansas 72701

Reynold Krueger
University of Illinois
Urbana, IL 61801

Albert Kurtz
Apt. 418
Park Knowles Apartments
Winter Park, FL 32789

W. L. Landrum
3 Chisolm Street
Charleston, South Carolina 29401

Rex Leonard
University of Southern Mississippi
Hattiesburg, Mississippi 39401

James Linden
Department of Psychology
Purdue University
Lafayette, Indiana 47907

Janice Lokan
5 Taunton Place
Ottawa K1J 7J7, Ontario
CANADA

Judy Lyon
CEMREL
10646 St. Charles Rock Rd.
St. Ann, MO 63074

Frank Many
2220 Piedmont Avenue
Berkeley, CA 94702

R. A. Martin
College of Education
University of Toronto
371 Bloor St. W.
Toronto, Ontario
CANADA

V. Rutledge McClaran
1722 W. Oak
Denton, TX 76201

John E. McClure
West Virginia Department of Education
Building 6B, Room 337
State Capitol
Charleston, West Virginia 25305

John McConnell
1415 South Wolf Road, #205
Wheeling, IL 60090

William McCormick
Bureau of Evaluation and Research
State Education Building
721 Capitol Mall
Sacramento, CA 95814

Garnet McDiarmid
Ontario Institute for Studies in Ed.
Toronto, Ontario
CANADA

Norman B. McEachron
Room Mo-201
Stanford Research Institute
333 Ravenswood Avenue
Menlo Park, CA 94025

Lawrence McNally
Board of Cooperating Educational Services (BOCES)
125 Jericho Turnpike
Jericho, NY 11753

Keith McNeil
Dept. of Guidance & Ed. Psych.
Southern Illinois University
Carbondale, IL 62901

W. V. Meredith
1231 S. E. 14th Court
Deerfield Beach, Florida 33441

Paul Merrill
Florida State University
1A Tully Building
Tallahassee, Florida 32306

Isadore Newman
Dept. of Educational Foundations
University of Akron
Akron, Ohio

Ronald Nuttall
Associate Professor
Institute of Human Sciences
Boston College
Chestnut Hill, MA 02167

George Olson
2325 West Pensacola
Apt. 103
Tallahassee, Florida 32304

Raleigh Pegram
Dallas Independent Schools
3700 Ross Avenue
Dallas, TX

Dan Perkuchin
Department of Sociology
Bowling Green State University
Bowling Green, Ohio 43403

Vincent J. Piraino
11791 Birchwood Lane
Franklin, WI 53132

John T. Pohlman
1915 Gartside
Murphysboro, IL 62966

Miss Marjorie Powell
20 Dartmouth Place
Boston, MA 02116

Thomas Pyle
Dept. of Psychology
Eastern Washington State College
Cheney, Washington 99004

Nambury Raju
301 Plainfield Road
LaGrange, IL 60525

Nicholas Rayder
Clairmont Hotel
Berkeley, CA

Cheryl L. Reed
Rt. 9, Box 72
W. Lafayette, Indiana

James Reynolds
Ritenoar Consolidated School District
2420 Woodson Road
Overland, MO 63114

Carol Ritter
Computer & Data Processing Center
Carter Hall
University of Northern Colorado
Greeley, CO 80631

Emmett Ritter
Associate Professor
Educational Administration
McKee 419
University of Northern Colorado
Greeley, CO 80631

Bruce Rogers
Measurement & Statistics
College of Education
University of Maryland
College Park, MD 20742

Bob Rosemier
Educational Administration & Services
Northern Illinois University
DeKalb, IL

Abby Rosenfield
Northeastern Illinois University
Chicago, IL

Nolan Russell
4507 Berkley Street
Harrisburg, PA 17109

Gerald Schluck
4037 North Monroe
Tallahassee, Florida 32301

Terry Schurr
Ball State University
Muncie, Indiana

Donald Senter
Research Director
Educational Developmental Laboratories, Inc.
Huntington, NY 11743

Ping Kee Siu
23 East 17th Street
Apt. 1C
Brooklyn, NY 11226

Vincent Silluzio
18 Fairfield Street
Newtonville, MA 02160

O. Suthern Sims, Jr.
Dean, Student Affairs
University of Georgia
Athens, GA 30601

Theodore Sjoding
15920 S. W. 84th Avenue
Miami, Florida 33157

Donald Smith
3009 Amherst Road
Muncie, Indiana 47304

John C. Soderstrum
P.O. Box 13677
University Station
Gainesville, Florida 32601

Steve Spaner
3474 Beldeer Drive
Woodcliff Estates
St. Charles, MO 63301

Fay Starr
Psychology & Psychological Services
Southern Illinois University
Edwardsville, IL 62024

Alan Stewart
Associate in Education
Research
State Education Department
Albany, NY 12224

Gary Stock
Assistant Professor
Candler Hall
University of Georgia
Athens, GA 30601

Eric Strohmeyer
Assistant Professor
Reid Hall
Montana State University
Bozeman, Montana 59715

David Suddick
University of Georgia
Testing & Evaluation Center
Athens, GA 30601

Steve Teglovic, Jr.
School of Business
University of Northern Colorado
Greeley, CO 80631

Jerome Thayer
Director of Testing & Research
Union College
Lincoln, Nebraska 68506

Donald Thomas
76 Loretta Avenue
Apt. 2
Fairborn, Ohio 45324

David Thor
Educational Developmental Laboratories
McGraw-Hill
284 East Pulaski Road
Huntington, NY 11743

Margaret Trikalsky
University of Northern Cokorado
Greeley, CO 80631

Norman Uhl
407 Landerwood Lane
Chapel Hill, North Carolina 27514

Michael Vitale
4551 Likini Street
Honolulu, Hawaii

Karen Vroegh
Institute for Juvenile Research
1140 S. Paulina Street
Chicago, IL 60612

Arthur Walker
1500 Wilkinson
Marquette, MI 49855

Joe Ward, Jr.
Southwest Educational Laboratory
167 East Arrowhead Drive
San Antonio, TX 78228

William Ware
Assistant Professor
College of Education
University of Florida
Gainesville, Florida 32601

Robert Warmbrod
208 Agricultural Administration Building
The Ohio State University
2120 Fyffe Road
Columbus, Ohio 43210

George Wasem
300 State Office Building
Office of the Superintendent of Public
Instruction
Research Section
Springfield, IL 62706

G. Leighton Wasen
50 Wedgewood Terrace
Springfield, IL 62702

Dr. Billy-Belle Weber
605 Washington Place
East St. Louis, IL 62205

Bill Webster
Research & Evaluation
Dallas Independent School District
3700 Ross Avenue
Dallas, TX 75204

Donald Wells
Psychology Department
University of Tennessee at Martin
Martin, Tennessee 38237

George D. White
1588 Mokulua Dr.
Kailua, Hawaii 96734

John D. Williams
Bureau of Educational Research
University of North Dakota
Grand Forks, ND 58201

Makonnen Yimer
University of Illinois
Urbana, IL 61801

Virginia Zachert
Rock House - Jacks Creek
Route 1, Box 28
Good Hope, GA 30641