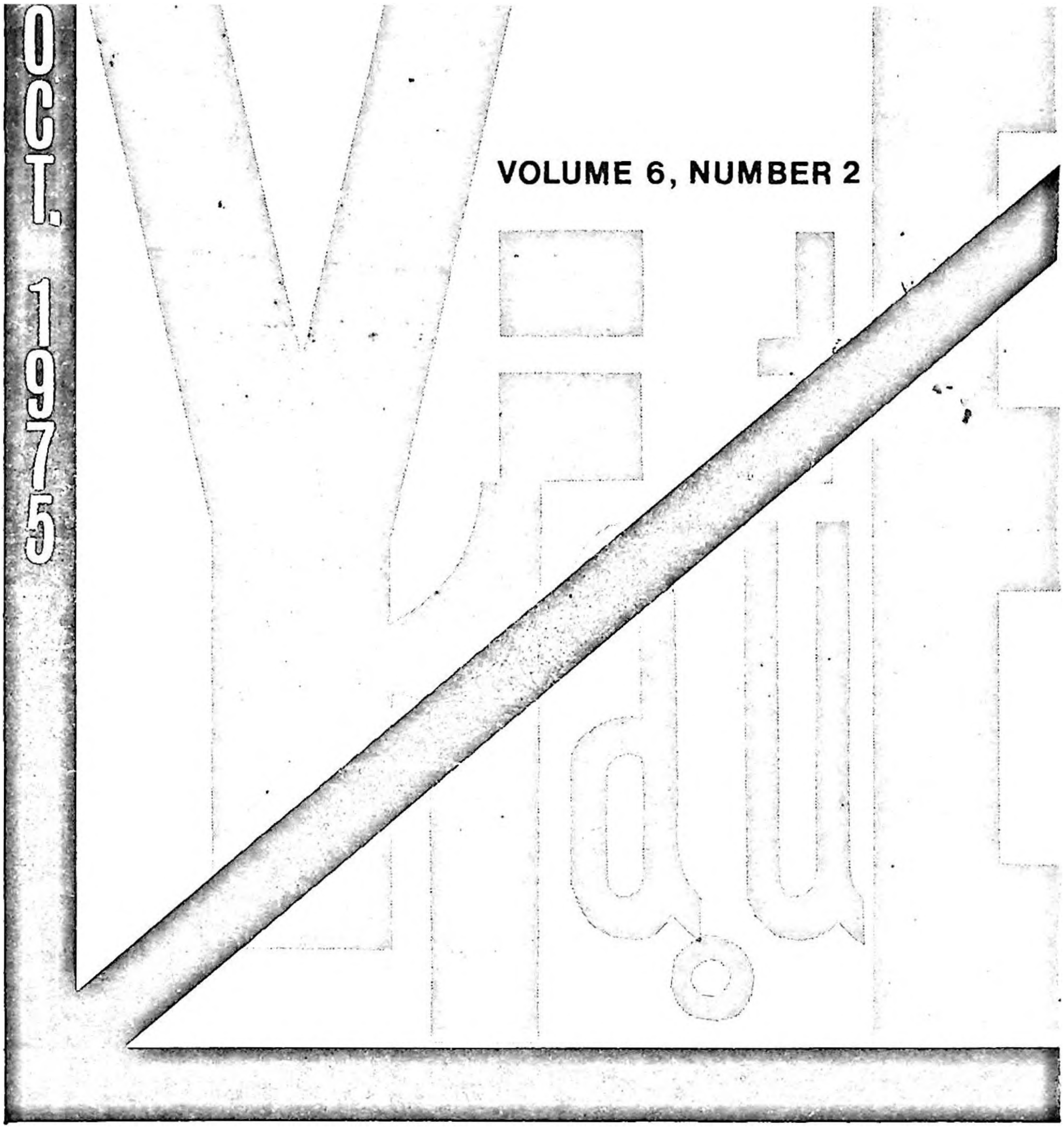

VOLUME 6, NUMBER 2

MULTIPLE LINEAR REGRESSION VIEWPOINTS

A publication of the Special Interest Group

on Multiple Linear Regression

MULTIPLE LINEAR REGRESSION VIEWPOINTS

TABLE OF CONTENTS

Multiple Linear Regression Viewpoints
Vol. 6, No. 2, 1975

THE RELATIONSHIP BETWEEN ACADEMIC PERFORMANCE, TEST ANXIETY, RACE, SEX, SCHOLASTIC ABILITY, AND SCHOOL ORGANIZATION: A MULTI-VARIABLE APPROACH

William Root, Isadore Newman & Edward Novak
The University of Akron

ABSTRACT

The relationships between academic performance, test anxiety, race, sex, scholastic ability, and school organization was investigated. It was found that the scholastic ability variable was the most predictive factor of academic performance. When covaring the scholastic ability variable, initial differences favoring Caucasian students in graded schools for academic performance and test anxiety became non-significant but significant differences between the sexes remained for for test anxiety. Caucasians, females, and those from graded schools scored significantly higher; however, when test anxiety was covaried, differences within school organizations became non-significant. Only linear significances were found, and all interactions were non-significant for the 206 students.

THEORETICAL FRAMEWORK

Anxiety, a twentieth century "discovery", has received much study since the origination of various anxiety scales during the past twenty-five years. Yet neither anxiety nor its relationship to performance (achievement) have reached a consensual status among researchers. (see Sarason, *et al.*, 1960, 1964; Ruebush, 1963; Castaneda, *et al.*, 1956; Allison, 1970; Deems, 1968; Lynn, 1957; and Taylor, 1956). Spielberger (1966) presented data showing anxiety research prior to 1950 as meager; however, since then over 1,500 studies have been listed in the *Psychological Abstracts* and more than 2,000 articles and books have appeared.

Some researchers attribute such increased activity to the formulation of various anxiety-measuring scales which stimulated the assessment of more than 120 personality-type tests to measure over 325 different anxiety variables by 1958. A popular and respected scale developed by S. B. Sarason entitled Test Anxiety Scale for Children (TASC) which measures anxiety in school within test-like situations was selected for use within this study.

Examination of selected literature established that the three dichotomous variables (race, sex, school organization) and the three continuous variables (academic performance, test anxiety, scholastic ability) were to be found within the research field, however, conclusive results as to their relational effects had been yet unrealized.

The examined school community of 23,000 inhabitants had an above-average level of schooling (college graduates > 30%), an above-average income production (mean income > $12,500), and a Caucasian population of 78% (U.S. Census Reports, 1972). The non-Caucasian income level surpassed its counterpart by more than $500 yearly (mean income) while a large random sampling of school parents exhibited an average mean income of $14,600 (Ross, 1971). Consequently, the importance of assessing the performance and characteristics of a racially mixed, middle-class socio-economic status suburban school system, with differing school organizational patterns, was viewed as significant since the researched literature had little to offer about samples of such composition.

METHOD AND DATA SOURCE

Sample: The Ss comprised the entire fourth grade/fifth year population of the three elementary schools within the district. There were 132 fourth graders and 74 nongraded fifth year students, 107 males and 99 females, 105 non-Caucasians and 101 Caucasians among the 206 students.

Procedure and Design: Standardized tests of scholastic ability and scholastic achievement were administered by the classroom teachers as normally performed so the testing could qualify as much as possible as a non-reactive measure and possess greater external validity. Appropriate levels of the Otis-Lennon Mental Abilities Test (OLMAT) and the Paragraph Meaning sub-test of the Stanford Achievement Test (PMT-SAT) were used. The Test Anxiety Scale for Children (TASC) was given to the ten classroom groups by the same person (Novak).

Multiple linear regression analysis was employed since it could accomodate for various interactions and curvilinear relationships while explicating the functional relationships among several relevant independent variables and a criterion behavior (Kelly, 1970). Academic performance (reading comprehension score) test anxiety and scholastic ability served as criterion variables while the test anxiety and scholastic ability variables also served as independent predictor variables with the dichotomous variables of race, sex, and school organization. Various interactions were examined as well as certain curvilinear relationships between the continous variables.

Significant results were obtained for sixteen of the thirty-four hypotheses.

The "set" of independent variables of test anxiety, race, sex, scholastic aptitude, and school organization accounted for a significant amount of variance for the criterion of academic (reading comprehension) performance (57% variance) with the scholastic aptitude variable being the best single predictor.

TABLE I

MODELS, F-RATIOS, AND R^2 FOR HYPOTHESIS I

DESCRIPTION OF MODELS	MODEL	R^2	df	F	p
Model$_1$ $Y_1 = a_oU + a_1x_1 \ldots$					
$+ a_8x_8 + E_1$	Full	.57	5/200	53.99	0.0***
Restriction: $a_1 = a_2 =$					
$a_3 = a_4 = a_5 = a_6 = a_7 = a_8 = \emptyset$					
Model $Y_1 = a_oU + E99$	Restricted	.00			

Hypothesis I:

An analysis of the series of independent variables selected (test anxiety, race, sex, scholastic ability, and school organization) accounts for a significant amount of the "explained" variance in the prediction of academic performance as represented by reading comprehension.

*** Significant beyond the .001 level ($p < .00001 < .05$)

See Table 5 and 6 for a description of all variables and models.

The continuous variables (academic performance, test anxiety, and scholastic aptitude or ability) were significantly related in linear fashion to one another: academic performance to scholastic ability (+.75); academic performance to test anxiety (-.28); and test anxiety to scholastic ability (-.38). The extent of curvilinearity between the continuous variables was not significant at alpha .05; thus, the relationships between the continuous variables were only linear in nature and did not lend support to the notion of a curvilinear relationship as stated by some researchers.

TABLE 2

MODELS, F-RATIOS AND R^2 FOR HYPOTHESIS IX

	DESCRIPTION OF MODELS	MODEL	R^2	df	F	p
Model 9	$Y_1 = a_oU + a_1x_1 +$					
	$a_{34}x_{34} + E9$	Full	.08			
	Restriction: $a_{34}=0$			1/203	0.9	.32
Model 8	$Y_1 = a_oU + a_1x_1 +$					
	E_8	Rest.	.07			

Hypothesis IX:

An analysis of the curvilinear relationship between the predictor variable, test anxiety, and the criterion of academic performance (reading comprehension) accounts for a significant amount of variance above and beyond the variance accounted for by the linear relationship between the two variables.

TABLE 3

MODELS, F-RATIOS AND R^2 FOR HYPOTHESIS XI

Model 11	$Y_1 = a_oU + a_6x_6 +$					
	$a_{35}x_{35} + E_{11}$	Full	.568			
	Restriction: $a_{35}=0$			1/203	1.7	.18
Model 10	$Y_1 = a_oU + a_6x_6 +$					
	E_{10}	Rest.	.565			

Hypothesis XI:

An analysis of the curvilinear relationship between the predictor variable, scholastic ability, and the criterion of academic (reading) performance accounts for a significant amount of variance above and beyond the variance accounted for by the linear relationship between the two variables.

Students from the self-contained, graded school organization and students of the Caucasian race had significantly higher mean scores for academic (reading comprehension) performance and significantly lower mean scores for test anxiety performance than their counterparts (nongraded organization and non-Caucasians); however, with the scholastic ability variable covaried, the mean score differences between these dichotomous groups became non-significant. There were no significant mean score differences between the sexes for any analyses made for academic performance, but females did score significantly higher mean scores for test anxiety with and without covariation of the scholastic ability variable.

Students within the self-contained, graded, school organization, students of the Caucasian race, and students of the female sex had significantly higher mean scores for the scholastic ability variable than their counterparts. When test anxiety was covaried, the mean score differences remained significant for scholastic ability between the races and between the sexes, results favoring Caucasians and females. A five-way interaction between the "set" of predictors (sex, race, school organization, scholastic ability, and test anxiety) and two-way interactions between race and scholastic ability, between school organization and scholastic ability and between sex and scholastic ability provided no significant increase in predictability beyond the additive variance generated by the "set" of predictors for the criterion of academic performance; therefore, "global" statements about the linear relationships were statistically supported and permitted the "main" effect to represent the relationship across the entire range of the predictors' scores.

Of the thirty-four hypotheses tested, sixteen produced significant results ranging from $p < .00001 < .05$. Eight of the rejected hypotheses "approached" significance ($p > .10 < .20$). Only nine hypotheses were strongly rejected ($p > .20$). (See Table 4 for a summary of R^2, F-Ratios, and significance levels for the 34 hypotheses.). Table 7 (inter correlations between selected variables) is also presented for the reader's analysis beyond the hypotheses contained herein.

RESULTS AND CONCLUSIONS

The findings of this investigation suggested that test anxiety served as an interfering variable to successful school achievement performance and may have served as an interfering variable to successful usage of scholastic ability capacity. The Response Interference Theory (Ruebush, 1960, 1963; Sarason, et al., 1960, 1964; Gifford and Marston, 1966, for example) received support from this study although those researchers who hypothesized a differential reaction across the predictor variable of test anxiety for the criterion of test performance were unsupported within these results.

Based on the findings, educators should consider the reduction of test anxiety within the classroom an effective means to aid students improve in academic achievement. Findings indicated that heightened test anxiety levels were possessed by the less scholastically able and by the more poorly achieving students. Whether the "typical" ability and achievement tests discriminate against non-Caucasians at every socio-economic level has served for extensive discussion within the educational field (see Shockley, 1972; Gage, 1972; Valentine, 1971; Rist, 1971, for example) without final resolution. Debate also continues about what a scholastic ability test exactly does measure; however a position was established within this study which skirted the controversy to some degree (see Novak (1973) for a description).

Females did score higher than males on scholastic ability and test anxiety but not on achievement which substantiates the Response Interference Theory. The extended implications of the female-oriented elementary school and the overt behavioral differences between the sexes becomes too complex (and somewhat peripheral) to pursue within this study (see Sarason, 1960, 1964; Hill and Sarason, 1966 for an extended discussion about sex differences).

Implications for the schools lead toward a notion that test anxiety can be found among all achievers, high and low, and teachers should remain aware of this. How the child views himself in relation to others can be frequently overlooked by the adult world. Educators also need to develop a "sensitivity" to dependency traits in children; teachers need to look beyond the material to be learned and focus upon the interaction between student, material, and teacher. "Learning to cope" within the environment needs to be incorporated within the evaluative structure of the school. *

* * * * * *

* A replication of this study was recently conducted following two years spent by the sample within a nongraded, open pod school as a collective "mass". The authors of this paper are now in the process of analyzing the new data to note any new characteristics or changes registered by the pupils after two years of "common" schooling. During the initial study the students attended three separate and differentially organized elementary schools.

TABLE 4

A SUMMARY OF R^2, F-RATIOS, AND SIGNIFICANCE LEVELS FOR THE THIRTY-FOUR HYPOTHESES

HYPOTHESES	(Models)	R^2_f	R^2_r	df	F	P	Significance at the .05 level (N = 206)
I.*	1 vs 99	.57	.00	5/200	53.99	.0	Significant
II.	1 vs 2	.57	.57	1/200	.27	.59	Non-Significant
III.	1-3	.57	.57	1/200	.01	.89	Non-Significant
IV.	1-4	.57	.57	1/200	1.57	.21	Non-Significant
V.	1-5	.57	.14	1/200	202.22	.0	Significant
VI.	1-6	.57	.57	1/200	1.71	.19	Non-Significant
VII.	1-7	.59	.57	8/192	1.49	.15	Non-Significant
VIII.	8-99	.07	.00	1/204	17.6	.00004	Significant
IX.*	8-9	.08	.07	1/203	.9	.32	Non-Significant
X.	10-99	.56	.00	1/203	263.99	.0	Significant
XI.*	10-11	.568	.565	1/203	1.7	.18	Non-Significant
XII.	12-99	.14	.00	1/204	34.5	.0	Significant
XIII.	12-13	.14	.14	1/203	.04	.83	Non-Significant
XIV.	14-99	.07	.00	1/204	15.8	.0001	Significant
XV.	10-15	.567	.565	1/203	.79	.37	Non-Significant
XVI.	16-99	.002	.00	1/204	.6	.43	Non-Significant
XVII.	10-17	.569	.565	1/203	1.78	.18	Non-Significant
XVIII.	18-99	.035	.00	1/204	7.55	.006	Significant
XIX.	10-19	.57	.565	1/203	2.62	.10	Non-Significant
XX.	20-99	.025	.00	1/204	5.32	.02	Significant
XXI.	12-21	.146	.144	1/203	.49	.48	Non-Significant
XXII.	22-99	.021	.00	1/204	4.5	.03	Significant
XXIII.	12-23	.187	.144	1/203	10.7	.001	Significant
XXIV.	24-99	.02	.00	1/204	4.58	.033	Significant
XXV.	12-25	.153	.144	1/203	1.98	.16	Non-Significant
XXVI.	26-99	.092	.00	1/204	20.9	.00001	Significant
XXVII.	27-28	.205	.144	1/203	15.6	.0001	Significant
XXVIII.	29-99	.023	.00	1/204	4.90	.02	Significant
XXIX.	28-30	.189	.144	1/203	11.19	.0009	Significant
XXX.	31-99	.023	.00	1/204	4.92	.02	Significant
XXXI.	28-32	.154	.144	1/203	2.31	.12	Non-Significant
XXXII.	1-33	.579	.574	2/198	1.18	.30	Non-Significant
XXXIII.	1-34	.581	.574	2/198	1.59	.20	Non-Significant
XXXIV.	1-35	.574	.574	2/198	.0	1.0	Non-Significant

* Listed in tabular form within the paper.

See Tables 5 and 6 for a listing of models and a description of the variables included.

TABLE 5

THE COMPLETE SERIES OF REGRESSION MODELS WHICH REFLECT THE EMPIRICALLY TESTED FUNCTIONAL RELATIONSHIPS

* Model 1	$Y_1 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + E_1$
Model 2	$Y_1 = a_oU + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + E_2$
Model 3	$Y_1 = a_oU + a_1x_1 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + E_3$
Model 4	$Y_1 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + a_6x_6 + a_7x_7 + a_8x_8 + E_4$
Model 5	$Y_1 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_7x_7 + a_8x_8 + E_5$
Model 6	$Y_1 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + E_6$
Model 7	$Y_1 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + a_{26}x_{26} + a_{27}x_{27} + a_{28}x_{28} + a_{29}x_{29} + a_{30}x_{30} + a_{31}x_{31} + a_{32}x_{32} + a_{33}x_{33} + E_7$
* Model 8	$Y_1 = a_oU + a_1x_1 + E_8$
* Model 9	$Y_1 = a_oU + a_1x_1 + a_{34}x_{34} + E_9$
* Model 10	$Y_1 = a_oU + a_6x_6 + E_{10}$
* Model 11	$Y_1 = a_oU + a_6x_6 + a_{35}x_{35} + E_{11}$
Model 12	$Y_2 = a_oU + a_6x_6 + E_{12}$
Model 13	$Y_2 = a_oU + a_6x_6 + a_{35}x_{35} + E_{13}$
Model 14	$Y_1 = a_oU + a_2x_2 + a_3x_3 + E_{14}$
Model 15	$Y_1 = a_oU + a_2x_2 + a_3x_3 + a_6x_6 + E_{15}$
Model 16	$Y_1 = a_oU + a_4x_4 + a_5x_5 + E_{16}$

Model 17 $Y_1 = a_oU + a_4x_4 + a_5x_5 + a_6x_6 + E_{17}$

Model 18 $Y_1 = a_oU + a_7x_7 + a_8x_8 + E_{18}$

Model 19 $Y_1 = a_oU + a_6x_6 + a_7x_7 + a_8x_8 + E_{19}$

Model 20 $Y_2 = a_oU + a_2x_2 + a_3x_3 + E_{20}$

Model 21 $Y_2 = a_oU + a_2x_2 + a_3x_3 + a_6x_6 + E_{21}$

Model 22 $Y_2 = a_oU + a_4x_4 + a_5x_5 + E_{22}$

Model 23 $Y_2 = a_oU + a_4x_4 + a_5x_5 + a_6x_6 + E_{23}$

Model 24 $Y_2 = a_oU + a_7x_7 + a_8x_8 + E_{24}$

Model 25 $Y_2 = a_oU + a_6x_6 + a_7x_7 + a_8x_8 + E_{25}$

Model 26 $Y_3 = a_oU + a_2x_2 + a_3x_3 + E_{26}$

Model 27 $Y_3 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + E_{27}$

Model 28 $Y_3 = a_oU + a_1x_1 + E_{28}$

Model 29 $Y_3 = a_oU + a_4x_4 + a_5x_5 + E_{29}$

Model 30 $Y_3 = a_oU + a_1x_1 + a_4x_4 + a_5x_5 + E_{30}$

Model 31 $Y_3 = a_oU + a_7x_7 + a_8x_8 + E_{31}$

Model 32 $Y_3 = a_oU + a_1x_1 + a_7x_7 + a_8x_8 + E_{32}$

Model 33 $Y_1 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6$
$+ a_7x_7 + a_8x_8 + a_9x_9 + a_{10}x_{10} + E_{33}$

Model 34 $Y_1 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6$
$+ a_7x_7 + a_8x_8 + a_{11}x_{11} + a_{12}x_{12} + E_{34}$

Model 35 $Y_1 = a_oU + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6$
$+ a_7x_7 + a_8x_8 + a_{36}x_{36} + a_{37}x_{37} + E_{35}$

* Model 99 All models listed as 99 develop a variance equal to zero.

---

* Developed in tabular form in this paper.

TABLE 6

AN EXPLANATION OF THE SYMBOLS USED IN MODEL CONSTRUCTION

Y_1 = criterion score, Stanford Achievement Test, Paragraph Meaning Sub-Test, (PMT-SAT) reading comprehension; a dependent variable

Y_2 = criterion score, Test Anxiety Scale for Children (TASC); a dependent variable

Y_3 = criterion score Otis Lennon Mental Abilities Test (OLMAT); a dependent variable

x_1 = test anxiety score (TASC), a continuous predictor

x_2 = Caucasian race is 1; zero otherwise

x_3 = Non-Caucasian race is 1; zero otherwise

x_4 = Male sex is 1; zero otherwise

x_5 = Female sex is 1; zero otherwise

x_6 = score on Otis-Lennon Mental Abilities Test (OLMAT) listed as an I.Q. ratio score, a continuous predictor

x_7 = Nongraded school member is 1; zero otherwise

x_8 = Graded school member is 1; zero otherwise

x_9 = Caucasian race member X OLMAT score interaction (x_2*x_6)

x_{10} = Non-Caucasian race member X OLMAT score interaction (x_3*x_6)

x_{11} = OLMAT SCORE X nongraded school member (x_6*x_7)

x_{12} = OLMAT score X graded school member (x_6*x_8)

x_{13} = Caucasian race X male (x_2*x_4)

x_{14} = Caucasian males X graded school (x_{13}*x_8)

x_{15} = Caucasian males X non-graded school (x_{13}*x_7)

x_{16} = Caucasian race X female (x_2*x_5)

x_{17} = Caucasian females X graded school (x_{16}*x_8)

x_{18} = Caucasian females X non-grd school (x_{16}*x_7)

x_{19} = Non-Caucasian race X male (x_3*x_4)

x_{20} = Non-Caucasian males X graded school (x_{19}*x_8)

x_{21} = Non-Caucasian males X non-graded school (x_{19}*x_7)

x_{22} = Non-Caucasian race X female (x_3*x_5)

x_{23} = Non-Caucasian females X graded school (x_{22}*x_8)

x_{24} = Non-Caucasian females X non-graded school (x_{22}*x_7)

x_{25} = Interaction between TASC X OLMAT (x_1*x_6)

x_{26} = Caucasian males, graded school X TASC-OLMAT (x_{14}*x_{25})

x_{27} = Caucasian males, non-graded X TASC-OLMAT (x_{15}*x_{25})

x_{28} = Caucasian females, graded X TASC-OLMAT (x_{17}*x_{25})

x_{29} = Caucasian females, nongraded X TASC-OLMAT (x_{18}*x_{25})

x_{30} = Non-Caucasian males, graded X TASC-OLMAT (x_{29}*x25)

x_{31} = Non-Caucasian males, nongraded X TASC-OLMAT (x_{21}*x_{25})

x_{32} = Non-Caucasian females, graded X TASC-OLMAT (x_{23}*x_{25})

x_{33} = Non-Caucasian females, nongraded X TASC-OLMAT (x_{24}*x_{25})

x_{34} = The square of the score on TASC ($x_1 * x_1 = x_1^2$)

x_{35} = The square of the score on OLMAT ($x_6 * x_6 = x_6^2$)

x_{36} = The interaction between the male sex and OLMAT Score (x_4*x_6)

x_{37} = The interaction between the female sex and OLMAT Score (x_5*x_6)

TABLE 7

CORRELATIONS BETWEEN CERTAIN SELECTED VARIABLES
OF POTENTIAL INTEREST*

VARIABLES	1	2	3	4	5	6	7	8	9	10	11	12	13
1		-.28	+.26	-.26	-.05	+.05	+.75	-.18	+.18	+.36	-.19	-.14	+.32
2			-.15	+.15	-.14	+.14	-.38	+.14	-.14	-.20	+.12	+.12	-.21
3				-1.0	+.03	-.03	+.30	-.46	+.46	+.98	-.98	-.45	+.50
4					-.03	+.03	-.30	+.46	-.46	-.19	+.98	+.45	-.50
5						-1.0	-.15	-.07	+.07	+.02	-.06	-.08	+.03
6							+.15	+.07	-.07	-.02	+.06	+.08	-.03
7								-.15	+.15	+.43	-.19	-.07	+.32
8									-1.0	-.46	+.46	+.18	-.97
9										+.46	-.46	-.18	+.97
10											-.97	-.45	+.53
11												+.46	-.49
12													-.96
13													

* Because of the confounding effect of the dichotomous variables' impact upon some correlation coefficients, only underlined correlations should be examined. Some dichotomous variables are correlated to other dichotomous variables which produce unusable correlational coefficients.

Variable 1 = Paragraph Meaning Test, Stanford Achievement Battery, Int. I
Variable 2 = Test Anxiety Scale for Children
Variable 3 = Caucasian Race
Variable 4 = Non-Caucasian Race
Variable 5 = Male Sex
Variable 6 = Female Sex
Variable 7 = Otis-Lennon Mental Abilities Test, Scholastic Ability
Variable 8 = Nongraded School Organization
Variable 9 = Graded School Organization
Variable 10 = Caucasian X Scholastic Ability
Variable 11 = Non-Caucasian X Scholastic Ability
Variable 12 = Nongraded Organization X Scholastic Ability
Variable 13 = Graded Organization X Scholastic Ability

Correlations of $\pm .20$ required for significance at the .05 corrected α_a

REFERENCES

Castaneda, A., Palermo, D.S., and McCandless, B.R. "Complex Learning and Performance as a Function of Anxiety in Children and Task Difficulty," Child Development, XXVII (1956), 327-332.

Dean, Walter, Denzler, L., and Sarason, I.G. "Anxiety and the Intellectual Performance of High School Students," Child Development, XXXV (1964), 917-926.

Gage, N.L. "Replies to Shockley, Page and Jensen: The Causes of Race Differences in I.Q.," Phi Delta Kappan, LIII, 7 (March, 1972), 422-427.

_______, "I.Q., Heritability, Race Differences, and Educational Research," Phi Delta Kappan, LIII, 5 (January, 1972).

Gifford, Edith M., and Marston, Albert R. "Test Anxiety, Reading Rate, and Task Experience," Journal of Educational Research, LIX (1966), 303-306.

Hill, Kennedy, and Sarason, Seymour B. The Relation of Test Anxiety and Defensiveness to Test and School Performance Over the Elementary School Years. Monographs of the Society for Research in Child Development. Chicago, Illinois: University of Chicago Press, 1966.

Kelly, Francis J. "Multivariate Design Considerations," in Phillip Emmert and William D. Brooks (eds.) Methods of Research in Communication. New York: Houghton-Mifflin, 1970.

Kelly, F.J., Beggs, D.L.,McNeil, K.A., Eichelberger, T., and Lyon, J. Multiple Regression Approach. Carbondale, Illinois: Southern Illinois University Press, 1969.

Lynn, R. "Tempermental Characteristics Related to Disparity of Attainment in Reading and Arithmetic," British Journal of Educational Psychology, XXVII (1957), 62-67.

Novak, Edward L. "The Relationship Between Academic Performance, Test Anxiety, Race, Sex, Scholastic Ability, and School Organization of Pre-Adolescent Public School Students: A Multivariable Approach." Unpublished Doctoral Dissertation, The University of Akron, Akron, Ohio, 1973.

Rist, Ray C. "Student Social Class and Teacher Expectations: The Self-Fulfilling Prophecy in Ghetto Education," in Harvard Educational Review Reprint Series, No. 5, Challenging the Myths: The Schools, the Blacks, and the Poor. Cambridge, Massachusetts: Harvard Educational Review, 1971.

Ross, Dennis W. "A Survey: The Attitudes of Parents Toward the Warrensville Heights City School System." Unpublished Master's thesis, John Carroll University, Cleveland, Ohio, 1971.

Ruebush, Britton, K., Jr. "Children's Behavior As a Function of Anxiety and Defensiveness." Unpublished Doctoral dissertation, Yale University, New Haven, Connecticut, 1960.

_______, "Anxiety," in Child Psychology. Sixty-Second Yearbook, Part I, National Society for the Study of Education. H.W. Stevenson (ed.). Chicago: University of Chicago Press, 1963.

Sarason, S.B., Hill, K.T., and Zimbardo, P.G. A Longitudinal Study of the Relation of Test Anxiety to Performance on Intelligence and Achievement Tests. Monographs of the Society for Research in Child Development Chicago: University of Chicago Press, 1964.

_______, Davidson, K.S., Lighthall, F.F., Waite, R.R., and Ruebush, B.K. Anxiety in Elementary School Children. New York: John Wiley & Sons, Inc., 1960.

Shockley, William. "A Debate Challenge: Geneticity is 80% for White Identical Twins' I.Q.'s," Phi Delta Kappan, III, 7 (March, 1972), 415-419.

_______, "Dysgenics, Geneticity, Raceology: A Challenge to the Intellectual Responsibility of Educators," Phi Delta Kappan, LIII, 5 (January, 1972).

Spielberger, Charles D. (ed.) "Theory and Research on Anxiety," in Anxiety and Behavior. New York: Academic Press, 1966.

United States Bureau of the Census. 1970 General Social and Economic Characteristics: Census of the Population. Final Report Pc (1)-C37 Ohio. Wash. D.C.: U.S. Government Printing Office, April, 1972.

Valentine, Charles A. "Deficit, Difference, and Bicultural Models of Afro-American Behavior," in Harvard Educational Review Reprint Series No. 5, Challenging the Myths: The Schools, the Blacks, and the Poor. Cambridge, Massachusetts: Harvard Educational Review, 1971.

Multiple Linear Regression Viewpoints
Vol. 6, No. 2, 1975

PARENTAL INVOLVEMENT IN THE EDUCATION OF THEIR CHILDREN

William L. Duff, Jr. & Samuel R. Houston
University of Northern Colorado

Introduction

Relevance, responsiveness and restoration have joined the traditional "Three R's in the lexicon of the modern educator. Schools, it is said, have become too large and have developed many of the suffocating characteristics of a self-serving bureaucracy; they are not relevant to society nor responsive to the needs of the people they serve. Many maintain that the best hope to reverse this process is to involve community people in the educational decision making process. Parents and others interested in education should be restored a role in decisions heretofore made exclusively by schoolmen.

During the past few years a great deal of effort and some money has been expended to accomplish this end. Progress has been made. Public interest has been stimulated and community involvement encouraged. Most applaud this development as beneficial to the institution of public education but more importantly to the children our schools serve. Most also agree, however, that in order for progress to continue along these lines, the principal participants in the educational process--parents, school, and community people--must reach some common perceptions regarding their individual roles in the educational decision making process.

During the last decade, a great deal of attention has been given to expanding the role of parents in educational decision making. In most cases, however, the specific roles that parents are to assume have not been adequately defined. This lack of specificity has frequently generated a loss of trust in school administration and conflict among those involved. Consequently, the Far West Laboratory, recognizing the need, has initiated a program to survey the perceptions of the various educational stakeholders with respect to parental roles in educational programs. The purpose of this paper is to present and analyze the data collected by the Far West Laboratory.

The paper is divided into three major sections. Section I begins with a description of the evaluation instrument and data collection procedures used by the Far West Laboratory. This section ends with a presentation of the basic questions to be analyzed in this study. In Section II, the writers describe the multiple regression techniques used by the writers to gain leverage on the research questions. In Section III the writers analyze Far West's data using the Harding/Johnson cluster grouping process.

SECTION I

The Instrument:

The data was collected by survey technique. Six hundred twenty-one educational stakeholders in various Far West target school districts were asked to complete the Far West Laboratory Parent/Content Questionnaire. This instrument measures the perception of parental role by various educational stakeholders. The objective is to identify differences in the perceptions of these groups with regard to the role they feel parents should play in the educational decision making process. A copy of this questionnaire is found in Appendix I of this study.

The reader will note (from Appendix I) that the questionnaire combines eleven "parent" phrases and twenty-two "content" phrases to yield 242 statements that describe various degrees of parental control in educational decisions. Parent phrases indicate levels of parental control; e.g., parents should care, parents do care, etc. Content phrases indicate the program areas where this control is perceived. For example, a "parent" phrase--"Parents do make decisions about"--and a "content" phrase--"who teaches their children"--would be combined to form one statement: "Parents do make decisions about who teaches their children." This parent phrase would be similarly combined with the remaining 21 content phrases. The process is repeated over each of the eleven parent phrases, yielding the 22 descriptive statements.

The Population:

The 621 original respondents represented a cross section of 36 educational roles and nine geographical areas. These stakeholders include teachers, parents, administrators and allied education roles; e.g., nurses, teacher aides, etc., from urban, rural, and suburban school districts. Appendix II contains a list of the educational roles and geographical areas included in the survey.

The Data:

After the data were collected, key-punched, and verified, it was further sorted and categorized. Observations with missing responses were eliminated and the roles (02--teacher aide, 03--teacher assistant, 04--associate teacher) were combined. This reduced the data set from 621 to 550 observations. The resulting frequency table, which categorizes the data by role and district as it was used in the final data runs, is found in Appendix III.

The Questions:

In addition to providing the writers of this paper with pre-collected data, Far West Laboratory also specified the questions they wished answered. In generic terms the laboratory wished to gain leverage on the following questions:

1. What are the perceptions of parental roles currently held by educational stakeholder groups?

2. If these stakeholder groups differ with regards to where these perceptions, between which groups do these differences lie?

Their intent was to identify potential conflict areas between educational stakeholder groups and ultimately to implement policy to resolve these differences.

The writers proposed two complimentary analytical approaches in order to attack these questions; a variant of multiple regression analysis and the Harding/Johnson cluster grouping technique. In the remaining paragraphs of this paper, the writers will present the analysis associated with each of these techniques.

SECTION II

THE REGRESSION MODELS

The response of each of the 550 educational stakeholders to each of the 242 questions in the Parent/Content Questionnaire were recorded as either a one (indicating agreement with the statement) or as a zero (indicating disagreement with that particular statement). Thus, the stakeholders response to each item in the questionnaire was coded in binary form. By categorizing the stakeholders into mutually exclusive collectively exhaustive groups, they too can be identified in a binary coding system. Using data coded in this way it is possible to specify an interactive multiple regression equation; group membership being used as the independent and item response as the dependent variables. The raw regression coefficient of these regression equations, specified in this way, can be interpreted as the conditional probability of a favorable response to a Parent/Content question, given the responding stakeholders group membership.

The first step in specifying these regression equations was to group the various educational stakeholders into appropriate mutually exclusive groups. This was necessary in order to build manageable parsimony in the resulting equations as well as to provide a sufficient number of observations within each classification cell. Moreover, it was recognized that the classification agreed upon must expose the data in such a way that the results could be used in policy planning and program evaluation.

After discussion with the project director at the Far Western Laboratory, it was decided that two basic regression models should

Coding--Urbanization Model:

Once the data was classified, an interaction coding scheme was devised for each of the regression models. In the urbanization model, two variables (X_2 and X_3) were designated district variables and two variables (X_4 and X_5) were designated role variables. Variable designation for the remaining role and district variables (Role 1, District 1) were assigned to the unit vector in order to avoid introducing four additional binary variables into the model. Table 3 shows the coding scheme used in the Urbanization Model.

TABLE 3
CODING SCHEME--URBANIZATION MODEL

District	Role	Binary Variables								
		Unit Vector	District Variables		Role Variables		Interaction Variables			
		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
1	1	1	0	0	0	0	0	0	0	0
1	2	1	0	0	1	0	0	0	0	0
1	3	1	0	0	0	1	0	0	0	0
2	1	1	1	0	0	0	0	0	0	0
2	2	1	1	0	1	0	1	0	0	0
2	3	1	1	0	0	1	0	1	0	0
3	1	1	0	1	0	0	0	0	0	0
3	2	1	0	1	1	0	0	0	1	0
3	3	1	0	1	0	1	0	0	0	1

Determining the Probabilities--Urbanization Model:

Using the coefficients associated with the independent variables in this model, it is possible to generate a predicted response for each question (criterion). This predicted score may then be interpreted as the probability that the criterion will be one (i.e., 1 = favorable response to a parent/content question) given group membership.

For example, using D_iR_j to indicate the various joint occurances of district and role group membership, FR to indicate favorable response, and B_i to indicate the regression coefficients, we can say:

$$
\begin{aligned}
P(FR/D_1R_1) &= B_1 \\
P(FR/D_1R_2) &= B_1 + B_4 \\
P(FR/D_1R_3) &= B_1 + B_5 \\
P(FR/D_2R_1) &= B_1 + B_2 \\
P(FR/D_2R_2) &= B_1 + B_2 + B_4 + B_6 \\
P(FR/D_2R_3) &= B_1 + B_2 + B_5 + B_7 \\
P(FR/D_3R_1) &= B_1 + B_3 \\
P(FR/D_3R_2) &= B_1 + B_3 + B_4 + B_8 \\
P(FR/D_3R_3) &= B_1 + B_3 + B_5 + B_9
\end{aligned}
$$

Moreover, defining N_{ij} as the number of observations in each district-role combination, we can compute the marginal probabilities for each district and role.

$$P(D_1) = \frac{1}{\Sigma N_{1,j}} \left[P(FR/D_1R_1) \cdot N_{1,1} + P(FR/D_1R_2) \cdot N_{1,2} + P(FR/D_1R_3) \cdot N_{1,3}\right]$$

$$P(D_2) = \frac{1}{\Sigma N_{2,j}} \left[P(FR/D_2R_1) \cdot N_{2,1} + P(FR/D_2R_2) \cdot N_{2,2} + P(FR/D_2R_3) \cdot N_{2,3}\right]$$

$$P(D_3) = \frac{1}{\Sigma N_{3,j}} \left[P(FR/D_3R_1) \cdot N_{3,1} + P(FR/D_3R_2) \cdot N_{3,2} + P(FR/D_3R_3) \cdot N_{3,3}\right]$$

$$P(R_1) = \frac{1}{\Sigma N_{i,1}} \left[P(FR/D_1R_1) \cdot N_{1,1} + P(FR/D_2R_1) \cdot N_{2,1} + P(FR/D_3R_1) \cdot N_{3,1}\right]$$

$$P(R_2) = \frac{1}{\Sigma N_{i,2}} \left[P(FR/D_1R_2) \cdot N_{1,2} + P(FR/D_2R_2) \cdot N_{2,2} + P(FR/D_3R_2) \cdot N_{3,2}\right]$$

$$P(R_3) = \frac{1}{\Sigma N_{i,3}} \left[P(FR/D_1R_3) \cdot N_{1,3} + P(FR/D_2R_3) \cdot N_{2,3} + P(FR/D_3R_3) \cdot N_{3,3}\right]$$

The probability of a favorable response is simply the mean of the criterion.

<u>Coding--Implementation Model</u>:

The Implementation Model was coded in a similar fashion. It has two districts, three roles, and six interaction terms, plus the unit vector.

TABLE 4

CODING SCHEME--IMPLEMENTATION MODEL

		Binary Variables											
District	Role	Unit Vector X_1	District X_2	X_3	Role X_4	X_5	X_6	Interaction X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
1	1	1	0	0	0	0	0	0	0	0	0	0	0
1	2	1	0	0	1	0	0	0	0	0	0	0	0
1	3	1	0	0	0	1	0	0	0	0	0	0	0
1	4	1	0	0	0	0	1	0	0	0	0	0	0
2	1	1	1	0	0	0	0	0	0	0	0	0	0
2	2	1	1	0	1	0	0	1	0	0	0	0	0
2	3	1	1	0	0	1	0	0	1	0	0	0	0
2	4	1	1	0	0	0	1	0	0	1	0	0	0
3	1	1	0	1	0	0	0	0	0	0	0	0	0
3	2	1	0	1	1	0	0	0	0	0	1	0	0
3	3	1	0	1	0	1	0	0	0	0	0	1	0
3	4	1	0	1	0	0	1	0	0	0	0	0	1

The conditional and marginal probabilities for this model are computed in the same manner as in the Urbanization Model.

Determining Probabilities--An Example:

Before presenting the results of the actual regression runs, it might be useful to demonstrate the process used with a simple example. Table 5 shows a data listing for a three district, four role interaction model. This, of course, is the same classification scheme used in the Implementation Model. The reader should note there are an unequal number of observations in each district role classification (see Column 2). Column Three indicates the number of individuals that agreed with the question. Column Four indicates the proportion of the total number of people in each classification, that agreed with the example question. The remaining columns indicate the coding of the criterion (Column 5) and independent variables (Columns 6, 7, 8),

Three District, Four Role, Interaction Problem

1	2	3	4	5	6		7			8					
Identification Role & District	N	FR	$P(FR/D_1R_j)$	Criterion	District Variables		Role Variables			Interaction Variables					
					2	3	4	5	6	7	8	9	10	11	12
D_1R_1	3	2	2/3	1	0	0	0	0	0	0	0	0	0	0	0
				0	0	0	0	0	0	0	0	0	0	0	0
				1	0	0	0	0	0	0	0	0	0	0	0
D_1R_2	5	3	3/5	0	0	0	1	0	0	0	0	0	0	0	0
				0	0	0	1	0	0	0	0	0	0	0	0
				1	0	0	1	0	0	0	0	0	0	0	0
				1	0	0	1	0	0	0	0	0	0	0	0
				1	0	0	1	0	0	0	0	0	0	0	0
D_1R_3	10	5	5/10	0	0	0	0	1	0	0	0	0	0	0	0
				0	0	0	0	1	0	0	0	0	0	0	0
				0	0	0	0	1	0	0	0	0	0	0	0
				0	0	0	0	1	0	0	0	0	0	0	0
				0	0	0	0	1	0	0	0	0	0	0	0
				1	0	0	0	1	0	0	0	0	0	0	0
				1	0	0	0	1	0	0	0	0	0	0	0
				1	0	0	0	1	0	0	0	0	0	0	0
				1	0	0	0	1	0	0	0	0	0	0	0
				1	0	0	0	1	0	0	0	0	0	0	0
D_1R_4	3	2	2/3	1	0	0	0	0	1	0	0	0	0	0	0
				1	0	0	0	0	1	0	0	0	0	0	0
				0	0	0	0	0	1	0	0	0	0	0	0
D_2R_1	5	3	3/5	1	1	0	0	0	0	0	0	0	0	0	0
				1	1	0	0	0	0	0	0	0	0	0	0
				1	1	0	0	0	0	0	0	0	0	0	0
				0	1	0	0	0	0	0	0	0	0	0	0
				0	1	0	0	0	0	0	0	0	0	0	0

DATA LISTING (cont.)

1	2	3	4	5	6		7			8					
Identification Role & District	N	FR	$P(FR/D_iR_j)$	Criterion	District Variables		Role Variables			Interaction Variables					
					2	3	4	5	6	7	8	9	10	11	12
D_2R_2	10	4	4/10	0	1	0	1	0	0	1	0	0	0	0	0
				0	1	0	1	0	0	1	0	0	0	0	0
				0	1	0	1	0	0	1	0	0	0	0	0
				0	1	0	1	0	0	1	0	0	0	0	0
				0	1	0	1	0	0	1	0	0	0	0	0
				0	1	0	1	0	0	1	0	0	0	0	0
				1	1	0	1	0	0	1	0	0	0	0	0
				1	1	0	1	0	0	1	0	0	0	0	0
				1	1	0	1	0	0	1	0	0	0	0	0
				1	1	0	1	0	0	1	0	0	0	0	0
D_2R_3	5	3	3/5	1	1	0	0	1	0	0	1	0	0	0	0
				1	1	0	0	1	0	0	1	0	0	0	0
				1	1	0	0	1	0	0	1	0	0	0	0
				0	1	0	0	1	0	0	1	0	0	0	0
				0	1	0	0	1	0	0	1	0	0	0	0
D_2R_4	3	2	2/3	0	1	0	0	0	1	0	0	1	0	0	0
				1	1	0	0	0	1	0	0	1	0	0	0
				1	1	0	0	0	1	0	0	1	0	0	0
D_3R_1	5	3	3/5	0	0	1	0	0	0	0	0	0	0	0	0
				0	0	1	0	0	0	0	0	0	0	0	0
				1	0	1	0	0	0	0	0	0	0	0	0
				1	0	1	0	0	0	0	0	0	0	0	0
				1	0	1	0	0	0	0	0	0	0	0	0
D_3R_2	3	2	2/3	1	0	1	1	0	0	0	0	0	1	0	0
				1	0	1	1	0	0	0	0	0	1	0	0
				0	0	1	1	0	0	0	0	0	1	0	0

DATA LISTING (cont.)

1	2	3	4	5	6		7			8					
Identification Role & District	N	FR	$P(FR/D_iR_j)$	Criterion	District Variables		Role Variables			Interaction Variables					
					2	3	4	5	6	7	8	9	10	11	12
D_3R_3	5	4	4/5	1	0	1	0	1	0	0	0	0	0	1	0
				1	0	1	0	1	0	0	0	0	0	1	0
				1	0	1	0	1	0	0	0	0	0	1	0
				0	0	1	0	1	0	0	0	0	0	1	0
				1	0	1	0	1	0	0	0	0	0	1	0
D_3R_4	10	5	5/10	1	0	1	0	0	1	0	0	0	0	0	1
				1	0	1	0	0	1	0	0	0	0	0	1
				1	0	1	0	0	1	0	0	0	0	0	1
				1	0	1	0	0	1	0	0	0	0	0	1
				1	0	1	0	0	1	0	0	0	0	0	1
				0	0	1	0	0	1	0	0	0	0	0	1
				0	0	1	0	0	1	0	0	0	0	0	1
				0	0	1	0	0	1	0	0	0	0	0	1
				0	0	1	0	0	1	0	0	0	0	0	1
				0	0	1	0	0	1	0	0	0	0	0	1

Table 6 is a photo copy of the regression statistics resulting from running the data in Table 5 on a standard IBM multiple regression program.[1]

TABLE 6

MULTIPLE REGRESSION..... HOPESO

SELECTION..... 1

VARIABLE NO.	MEAN	STANDARD DEVIATION	CORRELATION X VS Y	REGRESSION COEFFICIENT	STD. ERROR OF REG. COEF.
2	0.34328	0.47839	-0.06629	-0.06666	0.38977
3	0.34328	0.47839	0.06060	-0.06665	0.38973
4	0.26866	0.44661	-0.08216	-0.06665	0.38976
5	0.29851	0.46106	0.04323	-0.16666	0.35133
6	0.23881	0.42957	-0.00527	0.00000	0.43577
7	0.14925	0.35903	-0.14132	-0.13333	0.48721
8	0.07463	0.26477	0.01882	0.16666	0.48721
9	0.04478	0.20837	0.04348	0.06665	0.59465
10	0.04478	0.20837	0.04348	0.13333	0.55121
11	0.07463	0.26477	0.13345	0.36665	0.48721
12	0.14925	0.35903	-0.05678	-0.10000	0.52474
DEPENDENT					
1	0.56716	0.49921			

INTERCEPT 0.66666

MULTIPLE CORRELATION 0.21792

STD. ERROR OF ESTIMATE 0.53371

By summing over the appropriate coefficients (see pages 6 and 7 above), we can see that:

TABLE 7

$$P(FR/D_1R_1) = (.66666) = .66666$$
$$P(FR/D_1R_2) = (.66666) + (-.06666) = .60000$$
$$P(FR/D_1R_3) = (.66666) + (-.16666) = .50000$$
$$P(FR/D_1R_4) = (.66666) + (\ .00000) = .66666$$
$$P(FR/D_2R_1) = (.66666) + (-.06666) = .60000$$
$$P(FR/D_2R_2) = (.66666) + (-.06666) + (-.06666) + (-.13333) = .40000$$
$$P(FR/D_2R_3) = (.66666) + (-.06666) + (-.16666) + (\ .16666) = .60000$$
$$P(FR/D_2R_4) = (.66666) + (-.06666) + (\ .00000) + (\ .06666) = .66666$$
$$P(FR/D_3R_1) = (.66666) + (-.06666) = .60000$$
$$P(FR/D_3R_2) = (.66666) + (-.06666) + (-.06666) + (\ .13333) = .66666$$
$$P(FR/D_3R_3) = (.66666) + (-.06666) + (-.16666) + (\ .36666) = .80000$$
$$P(FR/D_3R_4) = (.66666) + (-.06666) + (\ .00000) + (-.10000) = .50000$$

---

[1] IBM Application Program, System/360 Scientific Subroutine (360 A-CM-03X) Version III.

Also, we find that:

$$P(FR/D_1) = \frac{1}{21}\left[(.66666)(\ 3)+(.60000)(\ 5)+(.50000)(10)+(.66666)(\ 3)\right] = .57143$$

$$P(FR/D_2) = \frac{1}{23}\left[(.60000)(\ 5)+(.40000)(10)+(.60000)(\ 5)+(.60000)(\ 3)\right] = .52174$$

$$P(FR/D_3) = \frac{1}{23}\left[(.60000)(\ 5)+(.66666)(\ 3)+(.80000)(\ 5)+(.50000)(10)\right] = .60876$$

$$P(FR/R_1) = \frac{1}{13}\left[(.66666)(\ 3)+(.60000)(\ 5)+(.60000)(\ 5)\right] = .61538$$

$$P(FR/R_2) = \frac{1}{18}\left[(.60000)(\ 5)+(.40000)(10)+(.66666)(\ 3)\right] = .50000$$

$$P(FR/R_3) = \frac{1}{20}\left[(.50000)(10)+(.60000)(\ 5)+(.80000)(\ 5)\right] = .55000$$

$$P(FR/R_4) = \frac{1}{16}\left[(.66666)(\ 3)+(.66666)(\ 3)+(.50000)(10)\right] = .56250$$

and

$P(FR) = .56716$

The reader will note from the data listing that these results are the same proportions we built into this example problem (see Column 4, Table 5)..

Data Analysis--The Regression Models:

In the tables that follow, the writers use the technique described above on each of the 242 questions in the Parent/Content Questionnaire. and display the probabilities that result.[2] Tables 8 and 10, for example, list the probabilities generated for the Urbanization and Implementation Models, respectively. In both of these tables, a row represents the probabilities generated from a single regression run.

---

[2]The probabilities that result from these analyses are presented in this paper. The regression printouts are found under separate cover on file with the Far West Laboratory.

The rows are grouped, displaying the twenty-two content phrases on a single page. . Each table runs eleven pages--one page for each of the eleven parent phrases in the questionnaire.[3]

Tables 9 and 11 are simply a reaggregation of the probabilities displayed in Tables 8 and 10. In Table 9, the Urbanization data is displayed by grouping the eleven parent phrases on each of twenty-two pages, each page representing the same content phrase.

The plus and minus signs on these four tables (i.e., Tables 8, 9, 10, and 11) indicate that that particular group's response was significantly different (at the .01 level) from the typical response to that question holding either content (Tables 8 and 10) or parent (Tables 9 and 11) phrase constant.

Tables 12 and 13 are summary tables for the Urbanization Model. Table 13 summarizes the responses of the various stakeholders on each of the eleven parent phrases. Table 12 summarizes the responses of the various stakeholders on each of the twenty-two content phrases. The numbers in the body of each of these tables can be interpreted as the probability of a favorable response to a particular parent (Table 12) or content (Table 13) phrase. Tables 14 and 15 are interpreted similarly for the Implementation Models.

The same analytical pattern was used to interpret the data in each of these tables. The results of the analysis are found in the summary that preceeds each table.

---

[3]Parent Phrase #5, "Parents are denied any say about," is a negative phrase. The probabilities reported in Tables 9 through 15 have been adjusted to account for this fact. All probabilities associated with this phrase are reported in these tables as one minus the probabilities generated using the original data.

In these summaries, the writers first looked at the column headed by P(FR). The magnitude of the probabilities listed in this column indicates the existence of strong negative (low probability) or positive (high probability) feelings with regards to particular items in the Parent/Content Questionnaire.

Next, the columns displaying the conditional probability of a favorable response, across roles holding districts constant, were searched. The writer first looked for a sign change across these probabilities. This would indicate a conflict in perceptions among stakeholder groups. Then, the probabilities were searched to identify those items that had at least a twenty point spread in probability values. This indicates a lesser degree of conflict in judgement across the various stakeholder roles.

Finally, the writers looked at the columns displaying the conditional probability of a favorable response across districts holding roles constant. Again, a sign change was taken as evidence of extreme differences in perception and a twenty point spread as a minor difference.

Multiple Linear Regression Viewpoints
Vol. 6, No. 2, 1975

THE ANALYSIS OF INCOMPLETE DATA USING REGRESSION

Donald C. Weber
Miami University

1. Introduction

The purpose of this paper is to demonstrate how multiple regression analysis can conveniently be used to analyze "incomplete" experimental design data. This approach is particularly attractive today because of the wide-spread availability of computers and the arrival of time-sharing computer terminals on the scene.

Basically, the regression approach to the analysis of an experimental design calls for the reparametrization of the traditional design model to one of full rank. Mendenhall [4] introduces his readers to this reparametrized approach but later abandons it in favor of the usual sums of squares calculating procedures since he confines his presentation to complete and/or balanced data. Other authors ([2], [5], [6] and [7]) develop the general linear hypothesis which is fundamental to the regression approach but fail to exploit it in the sense of this paper. In [3] the regression approach is the central theme but the book lacks in underlying theory. First let us review the pertinent concepts needed for the examples to follow.

Consider the multiple regression model

$$Y_i = \beta_0 + \beta_1 x_i + \cdots + \beta_k x_{ki} + \varepsilon_i \quad , i = 1,2,\cdots,n \tag{1}$$

where

$$E(\varepsilon_i) = 0 \qquad \text{for all } i$$

$$E(\varepsilon_i \varepsilon_j) = \begin{cases} \sigma^2 & \text{if } i = j \text{ ; } i,j = 1,2,\cdots,n \\ 0 & \text{if } i \neq j \text{ ; } i,j = 1,2,\cdots,n \end{cases}$$

*Presented at the 1972 Joint Statistical Meetings, Montreal, Canada, August 14 - 17

In matrix notation we write

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \tag{2}$$

where $\underline{Y}$ and $\underline{\varepsilon}$ are $n \times 1$ random vectors, X is an $n \times (k + 1)$ matrix of known fixed quantities of rank $k + 1$, and $\underline{\beta}$ is a $(k + 1) \times 1$ vector of unknown parameters such that

$$E(\underline{\varepsilon}) = \underline{0} \quad \text{and} \quad \mathrm{Cov}(\underline{\varepsilon}) = \sigma^2 I ,$$

I being the $n \times n$ identity matrix. The corresponding sample form is written as

$$\underline{y} = X\underline{b} + \underline{e}$$

where the vector $\underline{b}$ is the estimate of vector $\underline{\beta}$ in (2).

Let $\underline{\beta}^*$ be a $q \times 1$ vector whose elements consist of a subset of the elements of $\underline{\beta}$ and let $\underline{0}$ be the $q \times 1$ zero vector. If we assume that $\underline{\varepsilon}$ is multivariate normal, then the hypothesis

$$H_0 : \underline{\beta}^* = \underline{0} \tag{3}$$

can be tested using the statistic

$$F = \frac{(SSE^* - SSE)/q}{SSE/(n - k - 1)} \tag{4}$$

which is a value of an F random variable with q and $n - k - 1$ degrees of freedom. Here SSE denotes the error sum of squares obtained fitting the "full" model and SSE^* denotes the error sum of squares corresponding to the "reduced" or "conditional" model which results when the condition $\underline{\beta}^* = \underline{0}$ is imposed on the full model. The procedure is demonstrated in the following example.

Consider the linear regression model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$$

which we shall call the full model, and suppose we wish to test the hypothesis

$$H_0 : \beta_2 = \beta_3 = \beta_4 = 0 .$$

Imposing H_0 on the full model, we arrive at the conditional model

$$Y = \beta_0^* + \beta_1^* x_1 + \epsilon^* .$$

For this example, the F-statistic (4) has 3 and n - 5 degrees of freedom and the quantity SSE^* - SSE is sometimes said to be the reduction in error sum of squares due to the inclusion of β_2, β_3 and β_4 in the model. Therefore some authors (e.g., Searle [7]) denote this reduction by

$$R(\beta_2,\beta_3,\beta_4|\beta_0,\beta_1) = SSE^* - SSE .$$

It should also be noted here that the regression sum of squares, denoted by SSR, is a quantity that, like SSE, is included in the output of any standard multiple regression computer program. Alternately we could compute $R(\beta_2,\beta_3,\beta_4|\beta_0,\beta_1)$ in the above example as

$$R(\beta_2,\beta_3,\beta_4|\beta_0,\beta_1) = SSR - SSR^* .$$

From least squares theory, t-statistics of interest available for the analysis of data using regression include

$$t = \frac{b_j - \beta_j}{s\sqrt{c_{jj}}} \quad , \quad j = 0,1,\cdots,k \tag{5}$$

where $s = \sqrt{SSE/(n - k - 1)}$ and c_{jj} is the j^{th} diagonal element of the inverse of the X transpose X matrix denoted by $(X'X)^{-1}$. This, of course, can be used to test a hypothesis or construct a confidence interval about a single parameter, β_j. Inferences in regard to a linear combination of the parameters can be made using the statistic

$$t = \frac{\underline{a}'\underline{b} - \underline{a}'\underline{\beta}}{s\sqrt{\underline{a}'(X'X)^{-1}\underline{a}}} \tag{6}$$

For example, a $1 - \alpha$ confidence interval for the E(Y) corresponding to a given $\underline{x}$ vector is

$$\underline{x}'\underline{b} \pm t_{\alpha/2}\, s\sqrt{\underline{x}'(X'X)^{-1}\underline{x}}\ . \tag{7}$$

The degrees of freedom associated with (5) and (6) is n - k - 1.

2. Analyzing a randomized incomplete block design

Searle [7] discusses the two-way crossed classification model without interaction in Section 1 of Chapter 7 in his book with reference to the data given in Table 1. If we look upon "stove brand" as a block and "make of pan" as a treatment the traditional representation of the model for the associated experiment is

$$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}\quad ,\quad i = 1,2,3,4\ ;\ j = 1,2,3\ . \tag{8}$$

Table 1. Randomized Incomplete Block Data

Brand of Stove	Make of Pan		
	A	B	C
X	18	12	24
Y	-	-	9
Z	3	-	15
W	6	3	18

Source: Searle [7], Table 7.1

Following Mendenhall [4], we can reparametrize this model to the regression model

$$Y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 + \alpha_5 x_5 + \varepsilon \tag{9}$$

where

$$x_1 = \begin{cases} 1, & \text{if Brand Y} \\ 0, & \text{otherwise} \end{cases}, \quad x_2 = \begin{cases} 1, & \text{if Brand Z} \\ 0, & \text{otherwise} \end{cases}, \quad x_3 = \begin{cases} 1, & \text{if Brand W} \\ 0, & \text{otherwise} \end{cases}$$

$$x_4 = \begin{cases} 1, \text{ if Pan B} \\ 0, \text{ otherwise} \end{cases}, \qquad x_5 = \begin{cases} 1, \text{ if Pan C} \\ 0, \text{ otherwise} \end{cases}$$

In terms of the parameters of the traditional model, it is easy to derive the correspondence $\alpha_0 = \mu + \beta_1 + \tau_1$ and

$$\alpha_1 = \beta_2 - \beta_1 \ , \quad \alpha_2 = \beta_3 - \beta_1 \ , \quad \alpha_3 = \beta_4 - \beta_1 \ ,$$

$$\alpha_4 = \tau_2 - \tau_1 \ , \quad \alpha_5 = \tau_3 - \tau_1 \ .$$

Hence the hypothesis that all treatment effects in model (8) are equal,

$$H_0 : \tau_1 = \tau_2 = \tau_3 \ ,$$

implies the hypothesis

$$H_0 : \alpha_4 = \alpha_5 = 0$$

with respect to model (9). Similarly, the hypothesis $\beta_1 = \beta_2 = \beta_3 = \beta_4$ implies $\alpha_1 = \alpha_2 = \alpha_3 = 0$. These hypotheses can then be tested by using test statistic (4) with a summary displayed in the usual analysis of variance table.

Using a multiple regression program we obtain the following results in analyzing the data in Table 1.

Full Model: $Y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 + \alpha_5 x_5 + \varepsilon$

$$SSE = 12$$

Conditional Model 1: $Y = \alpha_0^* + \alpha_1^* x_1 + \alpha_2^* x_2 + \alpha_3^* x_3 + \varepsilon^*$

$$SSE^* = 270 \text{ and } SSR^* = 162$$

Conditional Model 2: $Y = \alpha_0^{**} + \alpha_4^{**} x_4 + \alpha_5^{**} x_5 + \varepsilon^{**}$

$$SSE^{**} = 283.5 \text{ and } SSR^{**} = 148.5$$

The partial analysis of variance summarizing these numerical results appears in Table 2. Searle's R( ) notation is also given. Notice that in

Table 2. ANOVA for Randomized Incomplete Block Design

Source	d.f.	SS
Stove Brands	3	$271.5 = SSE^{**} - SSE = R(\beta\|\mu,\tau)$
Make of Pans	2	$258 = SSE^{*} - SSE = R(\tau\|\mu,\beta)$
Error	3	$12 = SSE$
Total	8	

this notation $SSR^{*} = R(\beta|\mu)$ and $SSR^{**} = R(\tau|\mu)$. In traditional statistical nomenclature, $R(\beta|\mu,\tau)$ and $R(\tau|\mu,\beta)$ are called the "adjusted" sums of squares while $R(\beta|\mu)$ and $R(\tau|\mu)$ are called the "unadjusted" sum of squares.

To illustrate some auxiliary statistical inferences, we obtain from the regression print-out the estimates

$$a_0 = 16,\quad a_1 = -17,\quad a_2 = -12,\quad a_3 = -9,\quad a_4 = -4,\quad a_5 = 10$$

and $s = 2$. Hence to compare Pan A with Pan C we test

$$H_0 : \tau_1 = \tau_3$$

and observe that $\tau_1 = \tau_3$ implies $\alpha_5 = 0$. Therefore we use (5) and calculate

$$t = \frac{a_5}{s\sqrt{c_{55}}} = \frac{10}{2\sqrt{2/3}} = 6.12 \text{ with 3 degrees of freedom.}$$

Next, suppose we wish to compare Pan C with the average of Pans A and B. The hypothesis

$$H_0 : 2\tau_3 = \tau_1 + \tau_2$$

implies $2\alpha_5 - \alpha_4 = 0$. Here we use (6) and calculate

$$t = \frac{\underline{d}'\underline{a}}{s\sqrt{\underline{d}'(X'X)^{-1}\underline{d}}} = \frac{24}{2\sqrt{2.25}} = 8 \text{ with 3 degrees of freedom}$$

where $\underline{d}' = [0 \quad 0 \quad 0 \quad 0 \quad -1 \quad 2]$. The elements of vector $\underline{a}$ are given above.

As an example of (7), let us obtain the 95% confidence interval for the mean response to Stove W and Pan C,

$$E(Y_{43}) = \mu + \beta_4 + \tau_3$$
$$= \alpha_0 + \alpha_3 + \alpha_5 \ .$$

We calculate

$$\underline{x}'\underline{a} \pm t_{.025}\, s\sqrt{\underline{x}'(X'X)^{-1}\underline{x}} = 17 \pm 4.86$$

where $t_{.025} = 3.182$ and $\underline{x}' = [1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1]$.

Parenthetically, it should be noted that the above inference calculations were easily carried out on a computer terminal using APL.

3. Analysis of an incomplete factorial

Bancroft [1] fits the factorial model

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \qquad (10).$$

$$i = 1,2,3,4,5 \ ; \quad j = 1,2 \ ; \quad k = 1,2,\cdots,n_{ij}$$

to the data found in Table 3. Because of the unequal subclass frequencies Bancroft spent several pages describing the traditional methods for obtaining the adjusted sums of squares for the analysis of variance. However, we shall see that the absence of orthogonality presents no complications whatsoever for the regression approach.

For our full model, we reparametrize model (10) to the regression model

$$Y = \gamma_0 + \gamma_1x_1 + \gamma_2x_2 + \gamma_3x_3 + \gamma_4x_4 + \gamma_5x_5 + \gamma_{15}x_1x_5 + \gamma_{35}x_3x_5 + \varepsilon \qquad (11)$$

where

$$x_1 = \begin{cases} 1, \text{ if } A_2 \\ 0, \text{ otherwise} \end{cases}, \quad x_2 = \begin{cases} 1, \text{ if } A_3 \\ 0, \text{ otherwise} \end{cases}, \quad x_3 = \begin{cases} 1, \text{ if } A_4 \\ 0, \text{ otherwise} \end{cases}$$

$$x_4 = \begin{cases} 1, \text{ if } A_5 \\ 0, \text{ otherwise} \end{cases}, \quad x_5 = \begin{cases} 1, \text{ if } B_2 \\ 0, \text{ otherwise} \end{cases} .$$

Table 3. Incomplete Factorial Data

A \ B	B_1	B_2
A_1	22, 25	-1, 40, 18
A_2	41, 41	23, 13
A_3	29, 29, 37	
A_4	49, 50	61
A_5	55	

Source: Bancroft [1], Table 1.10

Notice that the interaction terms enter into the model as products and that there are only two such terms (instead of four) because of the two empty cells. To obtain the partial analysis of variance as found in Table 4, we fit model (11) and the following conditional models to the data.

Conditional Model 1: $Y = \gamma_0^* + \gamma_1^* x_1 + \gamma_2^* x_2 + \gamma_3^* x_3 + \gamma_4^* x_4 + \gamma_5^* x_5 + \varepsilon^*$

Conditional Model 2: $Y = \gamma_0^{**} + \gamma_1^{**} x_1 + \gamma_2^{**} x_2 + \gamma_3^{**} x_3 + \gamma_4^{**} x_4 + \varepsilon^{**}$

Conditional Model 3: $Y = \gamma_0^{***} + \gamma_5^{***} x_5 + \varepsilon^{***}$.

Observe the anomaly that the main effects sums of squares are obtained by first assuming a no interaction model whereas the error sum of squares is calculated assuming interaction. This implies that the tests for main effects are unbiased only if interaction is not present. For example, it can be shown that $\gamma_5 = \beta_2 - \beta_1 + (\alpha\beta)_{12} - (\alpha\beta)_{11}$. Using the results displayed in Table 4, we find here that interaction is "not significant." As in the previous example, the regression sums of squares

$$SSR^{**} = 2572.30 = R(\alpha|\mu) \text{ and } SSR^{***} = 473.20 = R(\beta|\mu)$$

are referred to as unadjusted sums of squares in the literature.

Table 4. ANOVA for Incomplete Factorial

Source	d.f	SS
Treatments	7	$3213.77 = SSR = R(\alpha,\beta,\alpha\beta\|\mu)$
A(adjusted)	4	$2249.05 = SSE^{***} - SSE^{*} = R(\alpha\|\mu,\beta)$
B(adjusted)	1	$149.95 = SSE^{**} - SSE^{*} = R(\beta\|\mu,\alpha)$
AB	2	$491.51 = SSE^{*} - SSE = R(\alpha\beta\|\mu,\alpha,\beta)$
Error	8	$1041.67 = SSE$
Total	15	4255.44

Next, let us assume that it is appropriate to investigate the polynomial effects corresponding to factor A. Using traditional methods, the calculations for such effects are horrendous but again regression analysis does the job naturally and without strain. Although there are no additional complications when the levels of A are not equally spaced, for the purpose of illustration let us assume equal spacing. We begin with the model

$$Y = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_1^2 + \gamma_3 x_1^3 + \gamma_4 x_1^4 + \gamma_5 x_2 + \varepsilon$$

where

$$x_1 = \begin{cases} 0 \text{ , if } A_1 \\ 1 \text{ , if } A_2 \\ 2 \text{ , if } A_3 \\ 3 \text{ , if } A_4 \\ 4 \text{ , if } A_5 \end{cases} \qquad x_2 = \begin{cases} 1 \text{ , if } B_2 \\ 0 \text{ , if } B_1 \end{cases}$$

After fitting this polynomial model we successively fit the models

Conditional Model 1: $Y = \gamma_0^{(1)} + \gamma_1^{(1)} x_1 + \gamma_2^{(1)} x_1^2 + \gamma_3^{(1)} x_1^3 + \gamma_5^{(1)} x_2 + \varepsilon^{(1)}$

Conditional Model 2: $Y = \gamma_0^{(2)} + \gamma_1^{(2)} x_1 + \gamma_2^{(2)} x_1^2 + \gamma_5^{(2)} x_2 + \varepsilon^{(2)}$

Conditional Model 3: $Y = \gamma_0^{(3)} + \gamma_1^{(3)} x_1 + \gamma_5^{(3)} x_2 + \varepsilon^{(3)}$

Conditional Model 4: $Y = \gamma_0^{(4)} + \gamma_5^{(4)} x_2 + \varepsilon^{(4)}$.

The mechanics of partitioning factor A's sum of squares adjusted for factor B into its polynomial components is indicated in Table 5.

Table 5. Polynomial Effects for Factor A

Source	d.f.	SS
Linear	1	$1738.18 = SSR^{(3)} - SSR^{(4)}$
Quadratic	1	$66.25 = SSR^{(2)} - SSR^{(3)}$
Cubic	1	$20.14 = SSR^{(1)} - SSR^{(2)}$
Quartic	1	$424.49 = SSR - SSR^{(1)}$
Total (A adjusted)	4	2249.06

REFERENCES

[1] Bancroft, T. A. Topics in Intermediate Statistical Methods. Ames, Iowa: The Iowa State Press, 1968.

[2] Hoel, P. G., Port, S. C., and Stone, C. F. Introduction to Statistical Theory. Boston: Houghton Mifflin, 1971.

[3] Kelly, F. J., Beggs, D. L., and McNeil, K. A. Research Design in the Behavior Sciences: Multiple Regression Approach. Carbondale and Edwardsville: Southern Illinois University Press, 1969.

[4] Mendenhall, W. Introduction to Linear Models and the Design and Analysis of Experiments. Belmont, California: Wadsworth, 1968.

[5] Rao, C. R. Linear Statistical Inference and its Applications. New York: Wiley, 1965.

[6] Scheffe, H. The Analysis of Variance. New York: Wiley, 1959.

[7] Searle, S. R. Linear Models. New York: Wiley, 1971.

Multiple Linear Regression Viewpoints
Vol. 6, No. 2, 1975

ABSTRACTS FOR VOLUME 5

Richard Sutton
The University of Akron

Haynes, Jack R. and Swanson, Ronald G. Method For Comparison of Non-Independent Multiple Correlations. Viewpoints, 1975, Vol. 5, No.,1, 1-6.

Two types of Problems concerning the comparison of two or more non-independent correlations are presented. The first one concerns the ability to determine if the multiple correlation between a criterion and a set of predictors is significantly different from an R between the same criterion and a different set of predictors within the same sample. The second deals with the question of the differential predictive ability of a set of predictors from population to population. Rationale for a comparison proceedure is also given.

Maola, Joseph F. The Multiple Regression Approach For Analyzing Differential Treatment Effects - The Reversed Gestalt Model. Viewpoints, 1975, Vol. 5, No. 1, 1-6.

This paper presents a statistical procedure for examining differences in criteria above independent differences. The theoretical position underlying the approach is that "the sum of the parts is greater than the whole." A model is set up where a researcher is trying to determine whether the Gestalt method or the Behavioral method of group counseling is more effective for increasing the self concept of a certain population.

Byrne, John F. The Use of Regression Equations To Demonstrate Causality. Viewpoints, 1975, Vol. 5, No. 1, 11-22.

This article attempts to explain causality by bringing the reader an explanation of the meaning of R square (R^2) in a larger than mathematical sense, e.g., once it has been computed, what does it mean to the research project.

Pyle, Thomas W. Classical Analysis of Variance of Completely Within - Subject Factorial Designs Using Regression Techniques. Viewpoints, 1975, Vol. 5, No. 1, 23-32.

By presenting an example the author documents the standard ANOVA of completely within subject factorial design using regression analysis.

He concludes by making several points. (a) completely within subject factorial designs can be analyzed in the multiple regression frame work. (b) three different "full" models are necessary to obtain the correct F ratios. (c) it is particularly obvious that in order to obtain the correct F ratio for the A x B interaction one must assume that there is no A x B x S interaction.

Gloeckler, Theodore L. Use of Multiple Linear Regression in Analysis of Intelligence Test Score Changes for Visually Handicapped Adults. Viewpoints, 1975, Vol. 5, No. 1, 33-37.

Multiple linear regression was used to analyze complex data assessing longitudinal changes in I.Q. test performance of visually handicapped adults. Results indicated: (1) Patterns of performance similar to those found in sighted populations, and (2) no influence on I.Q. changes by a variety of ontological factors.

Newman, Isadore. Dinner with Dr. Earl Jennings. Viewpoints, 1975, Vol. 5, No. 1., 40-44.

Dr. Newman discusses his informal dinner conversation with Dr. Jennings at the last A.E.R.A. convention. The major part of the discussion centered around "shrinkage." Dr. Jennings stated: "Shrinkage is O.K. if you're interested in interpreting R^2." The article provides good fuel for thought.

Maola, Joseph F. Causality and Prediction: Similar But Not Synonymous. Viewpoints, 1974, Vol. 5, No. 2, 1-2.

A short cirtique of Byrne, (Viewpoints 1974, Vol. 5, No. 1, 11-22) and his article on causality.

Williams, John D. Regression Solutions to The AxBxS Design. Viewpoints, 1974, Vol. 5, No. 2, 3-9.

An alternative regression solution to the repeated measures design (AxBxS) is given, contrasting to a solution given earlier by Pyle (Viewpoints 1974, Vol. 5, No. 1, 23-32). The solution given here can be completed without altering the criterion measures to find the A and B effects. Also, the presenting solution can be translated quite easily into experimental design terminology.

St. Pierre, Robert S. Possible Relationships Between Predictor and Criterion Variables. Viewpoints, 1974, Vol. 5, No. 2, 10-27.

This article defines possible relationships between a set of predictors and a creterion, and presents a modified Vargen visual display and regression analysis appropriate to that relationship (sample $N = 1000$). Primary focus of article is on effect of each relationship on the square multiple correlation (R^2). Three cases are presented.

Houston, Samuel R. and Bolding, James T. Jr. Regression Chi-Square: Testing for a Linear Trend in Proportions in a 2 x c contingency table. Viewpoints, 1975, Vol. 5, No. 2, 28-31.

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.

Williams, John D. Four-Way Disproportionate Hierarchical Models. Viewpoints, 1974, Vol. 5, No. 2, 32-40.

A Solution for the four-way disproportionate fixed effects hierarchical model is given. Because there are 15 effects to be ordered, and the ordering will have a major impact on each different effect, the N - way hierarchical model should be used only when there is a strong *a priori* ordering suggesting itself to the researcher.

Houston, Samuel R. and Ohlson, E. LaMonte. Issues In Teaching Multiple Linear Regression For Behavioral Research. Viewpoints, Vol. 5, No. 2, 41-44.

Several questions, issues and approaches to teaching a basic course in Multiple Linear Regression (MLR) are raised in this article. These include the following: 1) MLR as a generalized procedure; 2) Use of matrix algelua in MLR; 3) Redundant models in MLR; 4) Orthogonal coding; and 5) data analysis and MLR. Suggestions and recommendations are made for the issues raised.

Deitchman, R.; Newman, I.; Burkholder, J.; and Sanders, R. E. An Application of the Higher Order Factorial Analysis Design with Disproportionality of Cells. Viewpoints, Vol. 5, No. 2, 45-57.

The results of the application of several techniques available for analyzing data when there is disproportionality of cells is reported. How accurately they answer the research questions asked is also discussed.

Williams, John D. and Klimpel, Ronald J. Path Analysis and Casual Models As Regression Techniques. Viewpoints, 1975, Vol. 5, No. 3, 1-20.

Using regression as a point of comparision path analysis and causal models get a detailed explanation. Land tenure and community leadership is an example used in the article to draw conclusions as to the analagous relationship between the recursive equations and the dropping of the corresponding partial regression coefficient.

Houston, Samuel R. and Ohlson, E. LaMonte. From Educational Evaluation to Decision Making: Jan to the Rescue. Viewpoints, 1975, Vol. 5, No. 3, 21-29.

Judgement analysis (JAN) and its application in several areas; general applications, evaluations, applications, evaluation of research proposals, and instructional program evaluation, were presented by the authors as a method for decision - capturing for decision makers. Examples of how the technique can be applied in each of the areas was presented.

Maola, Joseph F. and Weis, David M. Unique Variances od Creativity and Dogmatism for Predicting Counseling Success. Viewpoints, 1975, Vol. 5, No. 3, 30-35.

In this article the authors are comparing the results stepivise multiple regression of unique variances for predicting counseling success from creativity and dogmatism, and to describe the inique prediction of creativity and dogmatism for predicting counseling success. Twenty-three graduate students in guidance and counseling were administered the Torrance Test of Creative Thinking Figural B (1966) and the Rokeach D - scale (1960). Stepevise multiple regression was used to interprete the data. It was found that dogmatism accounted for 29 percent of the variance and creativity accounted for 11.1 percent of the variance. The analysis of unique variance found dogmatism accounting for 2.6 percent and creativity accounted for 3.7 percent of the variance.

Huston, Samuel R. and Bolding, James T., Jr. Part, Partial, and Multiple Correlation in Commonality Analysis of Multiple Linear Regression Models. Viewpoints, 1975, Vol. 5, No. 3, 36-40.

"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."

McNeil, Keith and McNeil, Judy. Some Thoughts on Continuous Interaction. Viewpoints, 1975, Vol. 5, No. 5, 41-46.

In this article the authors emphasize by example the importance of specifying your expectations of the interaction of your variables. You must specify interaction as linear or eurvilinear and which comlunations of the predictor variables will yield high/low criterion scores.

Edelburn, Carl E., and Ochsnes, David P. STWNULTR: A Computer Program to Expedite The Retrieval of Residual Scores. Viewpoints, 1975, Vol. 5, No. 5, 47-48.

Resedual gain analysis is described in general terms and a new computer program, STWMULTR, designed to retrieve and pucch residual scores was described. Input and output sample data cards are included.

George, Joseph D. Multiple Regression Techniques Applies to Test The Effect of Three Types of Special Class Placement on The Reading Achievement of Educable Mentally Retarded Pupils: Viewpoints, 1975, Vol. 5, No. 3, 51-70.

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 score 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.

Byrne J. F. and McNeil K. A. A Further Discussion of Issues Related to Causal Inference. Viewpoints, 1975, Vol. 5, No. 4, 1-10.

The role of causality in scientific investigation, and in particular as it relates to multiple regression needs to be further discussed in this journal. Statistical pragmatism (as measured by R^2) is seen as a necessary but not sufficient condition for ascertaining causality. A research strategy is described which will help the researcher define the level of causal inference to which he is entitled. Interpretations of R^2 will vary depending upon the level which the researcher has allowed.

Evaluation of Maola's criticisims (1974 B) of Byrne (1974) is also included.

Ross, Maryann and McCabe,. George P. A Comment on Pohlmann's Algorithm for Subject Selection in Multiple Regression Analysis. Viewpoints, 1975, Vol. 5, No. 4, 11-15.

Pohlmann's algorithm for incorporating cost criteria into the valuable selection problem in multiple regression is examined. It is pointed out that this algorithm has the property that the choice of the optimal subset can be artificially changed by the addition of another variable. An example is included to illustrate this property.

Burkholder, Joel; Kapusenski, David; and Deitchman, Robert. The Use of Regression Analysis in Rodent Maternal Behavior: Multiple Regression's Answer to the Age Confound. Viewpoints, 1975, Vol. 5, No. 4, 16-26.

This article presents the use of Regression Analysis as the most logical and meaningful way to analyze continuous developmental change. An experiment is presented using both a conventional analysis of variance for repeated measures and Multiple Linear Regression for purposes of comparing the two techniques.

McNeil, Keith A.; McNeil, Judy T.; and Kelly, Francis J. Another Viewpoint of Issues in Multiple Linear Regression. Viewpoints, 1975, Vol. 5, No. 4, 27-33.

Disagreement with Houston and Ohlson, (Viewpoints 1975, 5(4)) and there presentation of five issues in the teaching of MLR. The basis of the disagreement is the feeling by the authors that many statisticians fail to make destinctions between behavioral science researchers and statistics majors and the degree to which blind learning replaces true comprehension of the subject material.

Newman, Isadore and Newman, Carol. A Discussion of Component Analysis: Its Intended Purpose, Strengths, and Weaknesses. Viewpoints, 1975, Vol. 5, No. 4, 35-48.

An examination of commonality analysis, elements analysis or component analysis. This article was presented at the 1975 American Educational Research Association; Multiple Linear Regression Special Interest Group.

If you are submitting a research article 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-space 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.)

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.

"A publication of the Multiple Linear Regression Special Interest Group of the American Educational Research Association, published primarily to facilitate communication, authorship, creativity and exchange of ideas among the members of the group and others in the field. As such it is not sponsored by the American Educational Research Association nor necessarily bound by the Association's regulations.

THE UNIVERSITY OF AKRON
AKRON, OHIO 44325

AKRON OHIO
P.B. METER 337321
20

0004060 42 3
STEVE SPANER
BEHAVIORAL STUDIES AND RESEARCH
UN OF MO-ST LOUIS
ST LOUIS MO 63121

TABLE OF CONTENTS