VOLUME 13, NUMBER 1

SPRING 1984

MULTIPLE LINEAR REGRESSION VIEWPOINTS

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

MLRV is listed in EBSCO Librarians Handbook.
ISSN 0195-7171
Library of Congress Catalog Card #80-648729

TABLE OF CONTENTS

Volume 13, Number 1 Spring 1984

A Comparative Analysis of Books on Multiple Regression

Dennis W. Leitner

and

Randall E. Schumacker

Southern Illinois University at Carbondale

ABSTRACT

The purpose of this article was to compare 33 books on multiple regression that have appeared in past years. The books were compared on topics covered; year and publisher; orientation; level of presentation; background required; published reviews; and readability. This information should aid students and instructors in selecting a multiple regression book for a given topic, approach, or need.

INTRODUCTION

In the past decade and a half, many books and reference works on the topic of multiple regression have become available. This is probably due to a combination of several factors, some of which might be the following: (1) The widespread availability of computers and computer programs. (2) The pioneering work of Robert Bottenberg, Joe Ward, Earl Jennings and their disciples from Texas, and (3) The excellent articles by Jacob Cohen entitled "Multiple regression as a general data-analytic device" and Richard Darlington entitled "Multiple regression in psychological and research practice", both of which appeared in Psychological Bulletin. The books were compared on the following information: (1) topics covered; (2) year and publisher; (3) orientation (theoretical vs applied); (4) level of presentation (textbook vs reference); (5) background required (basic statistics vs matrix algebra); (6) reference list of published reviews; and (7) readability.

MULTIPLE LINEAR REGRESSION VIEWPOINTS

EDITORIAL BOARD

Ellinger (1981) presented a historical overview of readability formulae from 1921 to the present. She included the criteria on which they were based, factors evaluated, reliability and validity. The authors selected the Flesch Ease Formula (Flesch, 1948) because it is appropriate for adult reading material and does not require use of a word list. The Flesch Ease Formula is:

$$Score = 206.835 - (.846 * NSYL) - (1.015 * SL)$$

where: NSYL = number of syllables per a 100 word sample

SL = average sentence length

The reading ease score can range from 0 to 100 with a lower score indicating a more difficult reading level (see Table 3).

METHOD

Book Selection

The books selected for comparison represent those from the social sciences and statistical literature that an educational researcher might be inclined to use. The list is by no means exhaustive. The books are listed alphabetically by author in a separate section of the references.

Criteria for Topic Inclusion

The books were compared on selected topics that the authors felt readers would be interested in. A particular topic was indicated as being covered in a given book if sufficient presentation or explanation was present, e.g. an entry in the table of contents, subject index, or several pages of discussion. For example, linear regression meant the straight-line one predictor variable regression model.

Anova compared to regression implied a comparison of the two as special cases of the general linear model. Matrix algebra referred to whether or not the book contained matrix algebra computations or had an appendix with matrix algebra rules and procedures. The covariance topic included the discussion of analysis of covariance. The other selected topics are self-explanatory.

RESULTS

Table 1 indicates a comparison of topics among the books. Most books covered the following topics: linear regression, curvilinear, polynomial or nonlinear regression, zero-order correlation, matrix algebra, partial/semi-partial correlation, and anova compared to regression, respectively.

Several books published between 1966 and 1976 had included computer programs while later books did not because of the popularity of several statistical packages. For example, Younger (1979) published a book which included comparisons among SAS, SPSS, and BMDP statistical packages in performing numerous applied regression examples.

Table 2 indicates a comparison among the books according to year, publisher, orientation, presentation, and background. Most books appeared in the 1970's (1960's, n=9; 1970's, n=18; 1980's, n=6). The majority had an applied orientation (applied, n=21; theoretical, n=9; both, n=3) with both a textbook and reference level of presentation (textbook, n=9; reference, n=12; both, n=12). Most books also required a basic statistics background (basic statistics, n=21; matrix algebra, n=12). Books with a theoretical orientation usually required knowledge of matrix algebra.

Published reviews were found for many of the books. These are listed alphabetically by author in a separate section of the references. The reviews permit an individual to read about another person's opinion of a book the authors have reviewed.

Selected Topics by Author

Author Name	a	b	c	d	e	f	g	h	i	j	k	l
	Book Topics											
Belsley	x	x		x		x	x	x	x	x	x	x
Chatterjee	x	x					x	x		x	x	
Cohen (1)	x	x	x	x	x	x	x	x	x	x	x	x
Cohen (2)	x	x	x	x	x	x	x	x	x	x	x	x
Draper (1)	x	x	x		x	x		x	x	x	x	
Draper (2)	x	x	x		x	x	x	x	x	x	x	
Dunn	x	x	x		x	x						x
Edwards (1)	x	x			x	x						
Edwards (2)	x	x	x		x	x	x	x	x	x		x
Fraser	x			x					x			
Freund	x	x	x		x	x	x		x	x		x
Goldberger	x	x			x	x	x		x	x		
Graybill (1)	x	x	x		x				x			x
Graybill (2)	x	x	x	x	x	x			x	x	x	
Gunst	x	x			x		x	x	x	x	x	
Haitovsky	x								x			x
Huang	x	x		x	x		x	x	x	x		
Kelly	x	x	x	x	x	x						x
Kerlinger	x	x	x	x	x	x	x	x	x		x	x
Kleinbaum	x	x	x	x	x	x	x	x	x	x	x	x
Koerts	x				x				x	x		
Lewis		x	x			x		x				x
McNeil	x	x	x	x	x	x		x			x	x
Pedhazur	x	x	x	x	x	x	x	x	x	x	x	x
Plackett	x	x	x						x	x		
Rao	x		x		x	x			x			
Searle	x		x					x	x	x	x	x
Smillie	x	x			x			x	x	x	x	
Sprent	x	x	x		x	x			x	x		x
Ward	x	x			x			x	x			x
Williams	x	x	x		x	x		x			x	x
Wonnacott	x	x	x		x	x	x	x	x	x	x	x
Younger	x	x		x	x	x	x	x	x	x	x	

KEY: a = linear regression
b = curvilinear, polynomial, or nonlinear regression
c = anova compared to regression
d = multivariate techniques
e = zero-order correlation
f = partial/semi-partial correlation
g = multicollinearity
h = dummy, effect and/or contrast coding
i = matrix algebra
j = residual analysis/outliers
k = variable selection methods
l = covariance

Table 2

Comparative Book Information

Author	Year	Publisher	Orientation		Presentation		Background	
			a	b	c	d	e	f
Belsley	1980	John Wiley	x	x	x		x	
Chatterjee	1977	John Wiley	x			x		x
Cohen (1)	1975	John Wiley	x		x	x		x
Cohen (2)	1983	John Wiley	x		x	x		x
Draper (1)	1966	John Wiley	x		x	x		x
Draper (2)	1981	John Wiley	x		x	x	x	
Dunn	1974	John Wiley	x			x		x
Edwards (1)	1976	W.H.Freeman	x		x			x
Edwards (2)	1979	W.H.Freeman	x		x			x
Fraser	1979	McGraw-Hill		x		x	x	
Freund	1979	Marc-Dekker	x			x		x
Goldberger	1968	MacMillan		x		x	x	
Graybill (1)	1961	McGraw-Hill	x	x	x	x	x	
Graybill (2)	1976	Duxbury		x		x	x	
Gunst	1980	Marc-Dekker	x		x	x		x
Haitovsky	1973	Hafner		x		x	x	
Huang	1970	John Wiley		x		x	x	
Kelly	1969	SIU press	x		x			x
Kerlinger	1973	Holt,R & W	x		x	x		x
Kleinbaum	1978	Duxbury	x		x	x		x
Koerts	1969	Rotterdam	x	x	x		x	
Lewis	1978	SIU press	x			x		x
McNeil	1975	SIU press	x		x			x
Pedhazur	1982	Holt,R & W	x		x	x		x
Plackett	1960	Oxford press		x		x	x	
Rao	1965	John Wiley		x		x	x	
Searle	1971	John Wiley		x		x	x	
Smillie	1966	Ryerson	x		x			x
Sprent	1969	Methuen		x	x	x		x
Ward	1973	Prentice	x		x	x		x
Williams	1974	MSS corp.	x		x			x
Wonnacott	1981	John Wiley	x		x			x
Younger	1979	Duxbury	x		x	x		x

KEY: a = applied
b = theoretical
c = textbook
d = reference
e = matrix algebra
f = basic statistics

Table 3 indicates the Flesch Reading Ease Scale used to interpret the book score listed in Table 4. All books ranged from fairly difficult to very difficult which would be expected given the topics discussed. The readability measure does not take into consideration the numerous formulae, graphs, notation and mathematics. It does however provide some indication of readability for comparison among the books as well as a general indication of reading complexity compared to other types of reading material.

CONCLUSION

The information provided permits comparisons among several books of multiple regression published over past years. Certain topics were indicated as appearing in the majority of the books. Most of the books reviewed emphasized an applied orientation with a basic statistics background requirement. Additional inquiry about certain books is possible by referring to the published reviews. The Flesch Ease Formula was used to compute a score on each book. The books reflected a difficult reading level comparable to scientific and academic text.

Most books had an outstanding feature which became apparent during the review process. For example, Belsley covered analysis of outliers and sources of multicollinearity. Chatterjee covered multicollinearity, autocorrelation and ridge regression extremely well. Cohen (1) and Cohen (2) had the widest range of topics covered and included one of the few discussions of power. Draper (1), Draper (2), and Gunst present the analysis of residuals/outliers and variable selection techniques the best. Edwards (1) and Edwards (2) afford an excellent introduction to linear regression with the presentation of different designs for analysis with dummy, effect, and contrast coding. Graybill (1) and Graybill (2) offer a broad coverage of topics at an advanced level using a matrix algebra

Table 3

Flesch Reading Ease Scale[a]

Reading Ease Scale	Description	Typical Magazine	Grade Level
0 to 30	Very Difficult	Scientific	
30 to 50	Difficult	Academic	
50 to 60	Fairly Difficult	Quality	
60 to 70	Standard	Digests	8th
70 to 80	Fairly Easy	Slick-fiction	7th
80 to 90	Easy	Pulp-fiction	6th
90 to 100	Very Easy	Comics	5th
100			4th

[a]Adapted from Flesch, 1948, p 230

Table 4

Readability Comparisons

R.E. Score	Author	Description
16.5	Cohen (2)	
18.1	Pedhazur	Very Difficult
23.1	Belsley	
32.6	Kleinbaum	
32.9	Haitovsky	
33.1	Gunst	
33.8	Huang	
35.1	Cohen (1)	
35.9	Goldberger	
35.9	Smillie	
36.8	Draper (2)	
38.2	Fraser	Difficult
39.5	Kerlinger	
42.7	Rao	
45.4	Sprent	
46.2	Wonnacott	
46.5	Koerts	
48.5	Dunn	
48.8	Ward	
50.7	Draper (1)	
52.0	Chatterjee	
52.7	Lewis	
53.1	Searle	
54.4	Graybill (2)	
54.9	Freund	
55.9	Edwards (1)	Fairly Difficult
56.6	McNeil	
56.7	Williams	
56.9	Kelly	
57.2	Younger	
57.2	Graybill (1)	
57.4	Plackett	
60.0	Edwards (2)	

approach. Kelly, Lewis, and McNeil propose model formulation to test given research hypotheses. Kerlinger and Pedhazur cover dummy, effect, and contrast coding well. Pedhazur additionally included a computer program on LISREL. Kleinbaum provides a broad coverage of all topics with excellent multivariate examples. Williams provides excellent examples on coding repeated measure designs. And finally, Younger provides computer applications using SAS, BMDP, and SPSS. Overall, selection of a specific book for classroom use is in the "eyes of the beholder", but this information should permit an alternative to experimentation or chance selection.

REFERENCES

Books Reviewed

Belsley, D.A., Kuh, E., & Welsch, R.E. (1980). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: John Wiley & Sons.

Chatterjee, S., & Price, B. (1977). Regression Analysis by Example. New York: John Wiley & Sons.

Cohen, J., & Cohen, P. (1975). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. New York: John Wiley & Sons.

Cohen, J., & Cohen, P. (1983). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (2nd Ed.). New York: John Wiley & Sons.

Draper, N. R., & Smith, H. (1966). Applied Regression Analysis. New York: John Wiley & Sons.

Draper, N. R., & Smith, H. (1981). Applied Regression Analysis (2nd Ed.). New York: John Wiley & Sons.

Dunn, O. J., & Clark, V. A. (1974). Applied Statistics: Analysis of Variance and Regression. New York: John Wiley & Sons.

Edwards, A. L. (1979). An Introduction to Linear Regression and Correlation. San Francisco: W. H. Freeman.

Edwards, A. L. (1979). Multiple Regression and the Analysis of Variance and Covariance. San Francisco: W. H. Freeman.

Fraser, D. A. S. (1979). Inference and Linear Models. Great Britian: McGraw-Hill.

Freund, R. J., & Minton, P.O. (1979). Regression Methods. A Tool for Data Analysis. New York: Marcel Dekker.

Goldberger, A. S. (1968). Topics in Regression Analysis. New York: MacMillan.

Graybill, Franklin A. (1961). An Introduction to Linear Statistical Models (Vol. 1). New York: McGraw-Hill.

Graybill, Franklin A. (1976). Theory and Applications of the Linear Model. Massachusetts: Duxbury Press.

Gunst, R. F. & Mason, R. L. (1980). Regression Analysis and its Application: A Data-oriented Approach. New York: Marcel Dekker.

Haitovsky, Y. (1973). Regression Estimation from Grouped Observations. New York: Hafner Press.

Huang, D. S. (1970). Regression and Econometric Methods. New York: John Wiley & Sons.

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

Kerlinger, F. N., & Pedhazur, E. J. (1973). Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston.

Kleinbaum, D. G., & Kupper, L. L. (1978). Applied Regression Analysis and other Multivariable Methods. California: Duxbury Press.

Koerts, J., & Abrahamse, A. P. J. (1969). On the Theory and Application of the General Linear Models. Netherlands: Rotterdam University Press.

Lewis, E. L., & Mouw, J. T. (1978). The Use of Contrast Coefficients (Supplement to McNeil, Kelly, and McNeil, Testing Research Hypotheses Using Multiple Linear Regression). Illinois: Southern Illinois University Press.

McNeil, K. A., Kelly, F. J., & McNeil, J. T. (1975). Testing Research Hypotheses Using Multiple Linear Regression. Illinois: Southern Illinois University Press.

Pedhazur, E. J. (1982). Multiple Regression in Behavioral Research Explanation and Prediction (2nd Ed.). New York: Holt, Rinehart & Winston.

Plackett, R. L. (1960). Principles of Regression Analysis. London: Oxford University Press.

Rao, C. R. (1973). Linear Statistical Inference and its Application (2nd Ed.). New York: John Wiley & Sons.

Searle, S. R. (1971). Linear Models. New York: John Wiley & Sons.

Smillie, K. W. (1966). An Introduction to Regression and Correlation. Toronto: Ryerson Press.

Sprent, P. (1969). Models in Regression and Related Topics. London: Methuen.

Ward Jr., J. H., & Jennings, E. (1973). Introduction to Linear Models. New Jersey: Prentice-Hall.

Williams, J. D. (1974). Regression Analysis in Educational Research. New York: MSS Information Corporation.

Wonnacott, T. H., & Wonnacott, R. J. (1981). Regression: A Second Course in Statistics. New York: John Wiley & Sons.

Younger, M. (1979). A Handbook for Linear Regression. California: Duxbury Press.

Belsley

Patterson, H. D. (1981). Biometrics. 37. 862-863.

Chatterjee

Bargmann, R. E. (1979). Contemporary Psychology. 24 (1). 45-46.

Cohen (1)

Tufte, E. R. (1979). Journal of the American Statistical Assn. 74. 935.

Henry, Neil W. (1977). Contemporary Psychology. 6. 320.

Cohen (2): none found

Draper (1)

Mallows, C. L. (1967). American Statistical Association Journal. 62. 1075.

Draper (2): none found

Dunn

Imrey, P. B. (1976). Journal of American Statistical Association. 71. 245.

(1974). Choice. 7. 1170.

Edwards (1)

Timm, N. H. (1977). Contemporary Psychology. 22. 174.

Edwards (2)

Timm, N. H. (1980). Choice. 16. 1614.

Fraser: none found

Freund

Dwass, M. (1980). Science Books and Films. 16. 12.

Goldberger

Mandansky, A. (1970). American Statistical Association Journal. 64. 1965.

Graybill (1)

Cochran, W. G. (1962). American Statistical Association Journal. 57. 503.

Graybill (2)

Sen, P. K. (1978). American Statistical Association Journal. 73. 221.

Gunst

McGuigan, R. (1979). Science Books & Films. 15 (1). 194.

Haitovsky: none found

Huang

Geraci, V. J. (1972). American Statistical Association Journal. 67. 961.

Kelly
Wasik, J. L. (1971). Educational and Psychological Measurement. 31. 575-577.

Kerlinger
Elashoff, J. D. (1975). American Statistical Association Journal. 70 959.

(1974). Choice. 10. 1753.

Walsh, J. A. (1974). Educational and Psychological Measurement. 34. 724-728.

Kleinbaum
Bhattacharya, H. T. (1979). American Statistical Association Journal. 74. 732.

Koerts: none found

Lewis: none found

McNeil
Burstein, L. (1976). American Educational Research Association Journal. 13. 196-301.

Pedhazur: none found

Plackett
Cochran, W. G. (1961). American Statistical Association Journal. 56. 418.

Rao
Efron, B. (1967). American Statistical Association Journal. 62. 29

Searle
John, P. W. M. (1972). Journal of the American Statistical Association. 67. 954.

Smillie: none found

Sprent
(1970). Choice. 7. 717.

Ward
Burstein, L. (1976). American Education Research Journal. 13. 296-301.

Williams: none found

Wonnacott: none found

Younger
Askin, R. G. (1980). Journal of Quality Technology. 12. 55-56.

(1980). Choice. 17. 111.

Readability

Ellinger, P. J. (1981). Use of Readability Formulas in the Formative Evaluation of Patient Education Materials. Unpublished master's thesis, Southern Illinois University. Carbondale, Illinois.

Flesch, R. F. (1948) A new readability yardstick. Journal of Applied Psychology. 32. 221-233.

Articles

Cohen, J. (1968). Multiple regression as a general data-analytic system. Psychological Bulletin, 70, 426-443.

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

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Volume 13, Number 1 Spring 1984

MICROCOMPUTER-RESIDENT PROGRAM FOR THE ANALYSIS OF STRUCTURAL EQUATIONS

Lee M. Wolfle

Virginia Polytechnic Institute and State University

ABSTRACT

A microcomputer-resident program for the analysis of structural equations (PASE) has been designed to provide the causal modeler with all of the usually desired estimates of coefficients in recursive causal models. The program is interactive and self-documenting, and requires only a series of option selections by the user. Output includes all of the usual regression coefficients, plus total causal effects decomposed into direct and indirect causal effects.

Structural equation causal models provide a powerful aid to assist in the substantive interpretation of social and educational processes. Unlike straightforward regression analyses, structural equation analyses permit the measurement, not only of direct causal effects, but also of indirect causal effects through other, causally intervening independent variables (Finney, 1972). For example, it is now well understood that the primary reason father's occupational status is so closely associated with son's occupational status is not that sons directly inherit their father's status, but rather that sons of fathers with high status attain educations of a level that allow entry into occupations of higher status.

Wolfle (1980), among others, showed how application of the basic theorem (Duncan, 1966) or first law (Kenny, 1979) of path analysis could be used to aid in the interpretation of the causal effects of one variable in a model on another. While the application of the first law of path analysis provides a useful aid in interpretation, in many cases its computation is tedious in practice. Alwin and Hauser (1975), followed by Wolfle (1983), showed how a series of relatively simple regression equations could be used to estimate the direct and indirect causal effects in a hierarchical structural equation model.

The present paper describes a microcomputer program designed to provide the causal modeler with all of the usually desired estimates of coefficients in recursive causal models, and also yields all total, direct, and indirect causal effects implied by the hierarchical causal ordering of variables in the equation.

The microcomputer program accomplishes this goal by the simple expedient of requesting the user to supply information about the causal order of variables in the model, and with this information calculates a series of reduced-form equations from which the total causal effects of independent variables are stored in the computer's memory. The differences between the reduced-form coefficients, or total effects, and the fully specified, or direct effect, coefficients are the total indirect causal effects. Algebraic proofs of this relationship have been provided by Griliches and Mason (1972) and Wolfle (1983).

The microcomputer program described in this paper, PASE: Program for Analysis of Structural Equations, was designed to provide the causal modeler with all of the usually desired estimates of coefficients in recursive causal models. The program is user friendly in that it is interactive and self-documenting, and requires only a series of option selections by the user. The program requires the input of a zero-order correlation matrix from either an existing file or the keyboard. Output includes all of the usual regression coefficients, plus total causal effects decomposed into direct and indirect causal effects.

System Requirements

PASE was written for an Apple II or Apple II Plus microcomputer that utilizes Applesoft BASIC. System configuration must be a minimum of 48K RAM, and one disk drive operating on DOS 3.3. The program provides support for, but does not require, a printer for hard-copy output.

The Analysis of Structural Equations

The most important advance in social research methodology in the past 15 years has been the introduction (Duncan, 1966) to the social sciences of causal modeling techniques first worked out over 60 years ago (Wright, 1921, 1925). On the one hand, this development has been important to social theory, for the techniques of causal modeling provide an explicit link between theory and the equations used to test the hypothesized relationships. On the other hand, while the estimation methods for structural equations implied by causal models are not new, the techniques have proven to be invaluable aids in the interpretation of social data. One of the most important of these interpretative aids in causal modeling is the decomposition of zero-order associations among variables into various causal components (see Wolfle, 1980).

A zero-order association may develop for one or all of three reasons. The association may be spurious; that is, it can develop because two variables, say X and Y, are related because they are both caused by a prior variable, Z, or a set of Z's. To the extent that the relationship between X and Y is spurious, that portion is called a noncausal component of the zero-order association. The remaining portion of the association between X and Y is causal, and is called the total effect. Total effects may in turn be decomposed into direct causal effects and indirect causal effects. Direct effects in recursive causal models are nothing more than partial regression coefficients of a variable regressed on all causes of it. The indirect causal effect of one variable on another is that portion of the total effect that can be traced through causally intervening variables. Such coefficients, both direct and indirect, can be expressed in either standardized or unstandardized (metric) form; the latter are often preferred, because standardized coefficients are relatively unstable from sample to sample or across populations (Duncan, 1975; Kim and Mueller, 1976).

Users of structural equation techniques need to keep in mind, however, that the interpretations of causal effects are model specific. If the causal model is plausible, the variables within it credibly ordered and accurately measured, then the interpretations of effects within it are plausible. If these conditions are not met, however, then the interpretations based on faulty models are themselves faulty.

Program Input and Output

A new computer program written for the Apple microcomputer, called PASE (Wolfle, 1982), provides a potentially useful tool for estimating hierarchical, recursive causal models. Because such models depend upon least-squares estimation procedures, PASE provides all of the usual regression coefficients. In addition, PASE provides estimates of total causal effects, and decomposes these into direct and indirect components.

PASE permits the input of new correlation matrices along with means and standard deviations. All data matrices can be saved to disk for future analyses. The program thus permits either the input of new matrices or the reading of previously saved data. Data matrices can be reviewed, corrected, truncated, or expanded to the maximum-sized matrix (17 variables) analyzable with the 48K memory limits of the compiled version of PASE.

Once the data have been input, reviewed if desired, changed if necessary, and saved to disk as recommended, the program prompts the user for the number of equations in the causal model. The program next asks the user to specify the dependent variable, followed by a list of the independent variables. The program next requests the user to specify the causal order among the independent variables. With this information, the program proceeds with the calculation of all regression coefficients, both standardized and metric, and decomposes these into direct and indirect causal components. (If one desires, the noncausal component of an association may be calculated by the simple expedient of subtracting the total causal effect from the zero-order association.)

The output of PASE has been organized for easy review. The output menu gives the user the option of reviewing the regression results, the regression ANOVA table, the R-squares among the independent variables, and the decomposition of causal effects. If desired, all of these results may be directed to a printer.

The regression results include all metric slopes, beta weights, standard errors, and t-ratios for the independent variables. The value of the intercept and the R-square for the regression are also included.

The ANOVA table includes the usual regression, residual, and total sums of squares, along with their associated degrees of freedom and mean squares. From these the F-ratio is calculated, and presented along with the standard error of estimate and the regression R-square.

The R-squares among the independent variables may be viewed. A high value among these suggests the presence of multicollinearity, which if present causes regression coefficients to be unstable in the face of slight changes in the zero-order correlation coefficients (see Gordon, 1968). In addition, standard errors are often inflated, and highly correlated independent variables often (and implausibly) have regression coefficients of opposite sign (see, for example, Muffo and Coccari, 1982).

The table of causal decomposition presents the total effect of each independent variable, along with its direct effect and total indirect causal effect. If there are no intervening variables between the causal independent variablo and the caused dependent variable, then the total effect is the direct effect.

An Illustration

To illustrate the use of PASE, refer to the causal model illustrated in Figure 1. The model is based on some analyses presented in Duncan, Featherman, and Duncan (1972), and the data taken from Duncan (1968). Of particular interest in this model is the relationship between ability and earnings; what is tho expected relationship between intelligence and earnings, controlling for social background, educational training, and occupational prestige?

There are three endogenous variables in the model; therefore, there are three equations to be estimated. Focusing attention on the equation for earnings, X(1), one would specify upon request by the program that variable 1 is dependent. The user will then be asked to specify the variable numbers of the causes of X(1); therefore, the user will input variable numbers 2, 3, 4, 5, and 6, since all other variables in the model are hypothesized to cause earnings.

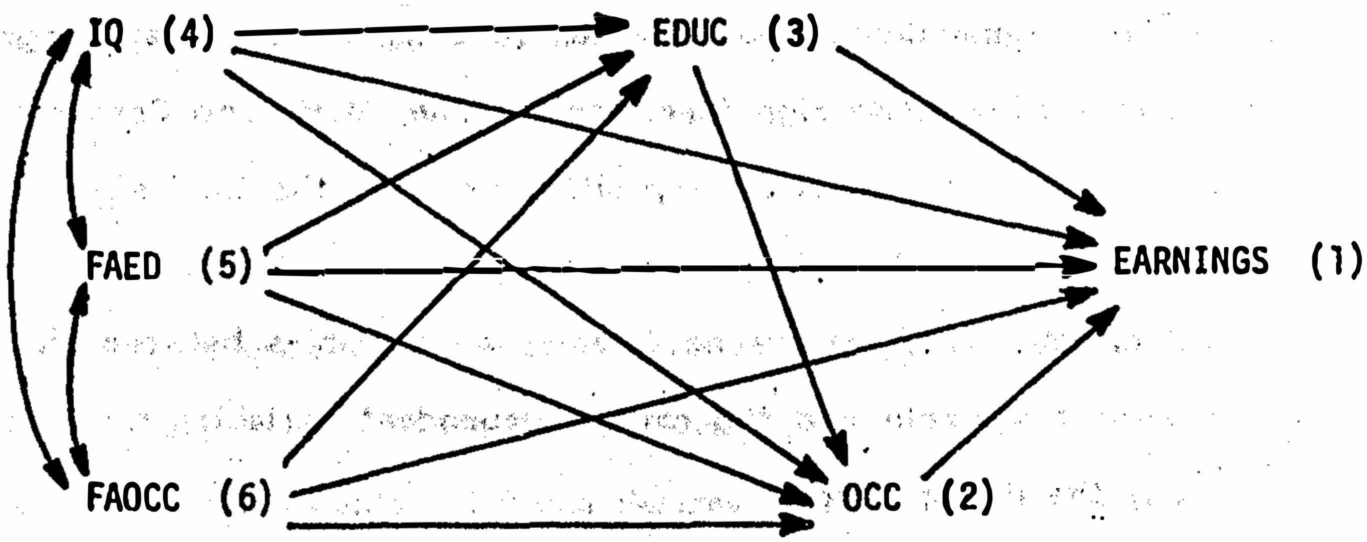

Figure 1. Path Diagram Representing Dependence of Earnings on Status Attainment, Intelligence, and Family Background.

The program will next ask the user to identify the causal order of the independent variables. In this case, X(4), X(5), and X(6) occur simultaneously in a single causal block of exogenous variables. The user would therefore input variable numbers 4, 5, and 6 as constituting the variables in Block 1 (followed by the value 99 to terminate the Block listing). Educational attainment, X(3), is the single variable in Block 2, followed by occupational prestige, X(2), in Block 3. With this blocking information, the program proceeds with the calculation of the coefficients for X(1).

The regression results for this equation are shown in Table 1. These indicate by examination of the beta weights that the most important effect of earnings is the prestige level of the respondent's occupation, X(2). The relative effects of education and intelligence are less than half that of prestige, while the influence of father's education and occupation are statistically indistinguishable from zero.

The decompositions of causal effects for this equation are shown in Table 2. Examination of the total causal effects indicate that intelligence, educational attainment, and occupational prestige all have about equal total effects on earnings. The indirect effects indicate that about half of the total effect of education on earnings occurs indirectly through occupation; that is, those people with higher levels of educational attainment not only receive higher earnings <u>ceteris paribus</u>, but also tend to enter occupations of higher prestige which in turn lead to higher earnings.

Table 1. Regression Results

Dependent Variable: 1

Var.	B	Beta	St.Err	T
2	.2625	.2625	.0381	6.8821
3	.1069	.1069	.0423	2.5277
4	.1013	.1013	.0344	2.9436
5	.0306	.0306	.0342	.8958
6	.0183	.0183	.0348	.5263

Variables:
1 = 1964 earnings,
2 = 1964 occupation,
3 = education,
4 = "early" intelligence,
5 = father's education,
6 = father's occupation.

Table 2. Decomposition of Causal Effects

(Standardized)

FROM	TOTAL	DIRECT	INDIRECT
VAR. 4	.2273	.1013	.1261
VAR. 5	.0881	.0306	.0574
VAR. 6	.1032	.0183	.0849
VAR. 3	.2454	.1069	.1385
VAR. 2	.2625	.2625	0

Variables: See Table 1.

The same may be said for the effects of intelligence on earnings. There is a direct causal effect of intelligence on earnings (the higher one's intelligence *ceteris paribus*, the higher one's earnings), but there is also a set of indirect effects wherein people of higher intelligence acquire higher levels of education and possess occupations of higher prestige, which also have positive effects on earnings. The combined direct and indirect effects of intelligence make it equally important to the explanation of earnings as is either education or occupational prestige.

In sum, PASE not only permits the causal modeler to examine the straightforward regression results, but, further, PASE also allows one to examine the decomposition of causal effects into their direct and indirect components. These latter examinations often prove to be very useful in revealing how causal effects are manifested in the model.

Availability of PASE

PASE is available from the author, College of Education, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061. Please enclose one blank 5.25-inch, soft-sectored floppy disk compatible with the Apple disk operating system. A users' guide is also available; to cover duplication costs, please enclose a check in the amount of $1.00 made out to VPI&SU College of Education.

REFERENCES

Alwin, Duano F., and Hauser, Robert M., "The Decomposition of Effects In Path Analysis." American Sociological Review, 1975, 40, 37-47.

Duncan, Otis Dudley, "Path Analysis: Sociological Examples." American Journal of Sociology, 1966, 72, 1-16.

Duncan, Otis Dudley, "Ability and Achievement." Eugenics Quarterly, 1968, 15, 1-11.

Duncan, Otis Dudley, Introduction to Structural Equation Models. New York: Academic Press, 1975.

Duncan, Otis Dudley, Featherman, David L., and Duncan, Beverly, Socioeconomic Background and Achievement. New York: Seminar Press, 1972.

Finney, John M., "Indirect Effects in Path Analysis." Sociological Methods & Research, 1972, 1, 175-186.

Gordon, Robert A., "Issues In Multiple Regression." American Journal of Sociology, 1968, 73, 592-616.

Griliches, Zvi, and Mason, William M., "Education, Income, and Ability." Journal of Political Economy, 1972, 80, S74-S103.

Kenny, David A., Correlation and Causality. New York: Wiley, 1979.

Kim, Jae-On, and Mueller, Charles W., "Standardized and Unstandardized Coefficients In Causal Analysis: An Expository Note." Sociological Methods & Research, 1976, 4, 423-438.

Muffo, John A., and Coccari, Ronald L., "Predictors of Outside Funding for Research Among AASCU Institutions." Research in Higher Education, 1982, 16, 71-80.

Wolfle, Lee M., "Strategies of Path Analysis." American Educational Research Journal, 1980, 17, 183-209.

Wolfle, Lee M., "PASE: Program for Analysis of Structural Equations." Behavior Research Methods & Instrumentation, 1982, 14, 548-550.

Wolfle, Lee M., "A Method for Estimating Indirect Effects In Path Analysis." Multiple Linear Regression Viewpoints, 1983 (in press).

Wright, Sewell, "Correlation and Causation." Journal of Agricultural Research, 1921, 20, 557-585.

Wright, Sewell, Corn and Hog Correlations. Washington, D.C.: U.S. Department of Agriculture Bulletin ~~1300~~, 1925.

COMPARISON OF PROCEDURES FOR TESTING THE HYPOTHESIS OF A DIFFERENCE BETWEEN r 1 AND r 2 USING INDEPENDENT AND DEPENDENT SAMPLES

Kenneth C. Hoedt, Isadore Newman and Gayle A. Seymour

The University of Akron

Introduction

Completion of correlation studies may require that the researcher test for significant differences between two independent correlations and/or between two dependent correlations. Solutions to the former problem may be found in many basic statistics books (Tate, 1965; McCall, 1970; Dayton, 1971; Minium, 1978). Procedures to test for a significant difference between dependent correlations have also been reported (Glass and Stanley, 1970; Hinkle, Wiersma and Jurs, 1979). Minium (1978) reported that there was no entirely satisfactory test of the difference between correlations from dependent samples, but it is not known whether he was familiar with the procedure presented by Hinkle, Wiersma and Jurs in 1979.

Method for Study One

Newman[1] suggested that differences between correlations from both dependent and independent universes could be tested for significance using multiple linear regression (MLR). This application of the use of MLR had not been previously demonstrated. While testing for a difference between r_s of independent universes appeared to be relatively uncomplicated using MLR, such was not the case when the test was applied to data from dependent universes. In the latter case repeated measurers were made, hence it was necessary to include Person Vectors in the statistical models developed. Peddhazur, 1977 reported a procedure for inclusion of Person Vectors in MLR models, but no analogue procedure was given when the dependent variable was dichotomous. This paper presents such an analogue procedure and demonstrates its appropriateness.

Results of using the procedure reported by Minium, 1978 to test for a significant difference between r_1 and r_2 using independent samples and the procedures reported by Glass and Stanley, 1970, and by Hinkle, Wiersma and Jurs, 1979 for testing the difference between r_1 and r_2 using dependent samples were compared to results using the general MLR approach suggested below. Study of the outcome for the independent sample case was based upon a Monte Carlo approach in which 100 pairs of samples of 30 subjects each were taken from the Coleman Data Bank. The criterion vari-

---

[1] Newman made the suggestion in planning the present paper.

able was sex (Y) and the predictor variables were GPA (X_1) and reading achievement (X_2). In the dependent case the same variables were used, but the subjects in sample 1 were the same subjects as those in sample 2. Using a Monte Carlo procedure, 100 samples of 60 subjects each were created from the Coleman Data Bank. When these subjects were considered to be in sample 1, a correlation (r_1) was calculated between GPA and sex. When the same subjects were in sample 2 a correlation (r_2) was calculated between reading achievement and sex.

Comparison of Minium's Suggestion (z test) to MLR for Testing $H_0: r_1 - r_2 = 0$, $H_A: r_1 - r_2 \neq 0$, $\alpha = .05$ for Independent Sample Data.

Using a Monte Carlo procedure 100 pairs of independent samples were drawn. Correlations (r_1 and r_2) were run between sex (Y) and GPA (X_1) and sex (Y) and reading achievement (X_2). To determine if there was a significant difference between r_1 and r_2 using the z test the following formula was applied:

Formula One:

$$z = \frac{z_{r_1} - z_{r_2}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

Fisher's z equivalents were used rather than r values because the sampling distribution of the r values is likely to be skewed. Values of z obtained for the 100 pairs of samples are reported in Table 1. Inspection of the z scores indicates that only two reached a magnitude greater than 1.96. Twice the null hypothesis was rejected with alpha set at the .05 level.

To determine if there was a significant difference between r_1 and r_2 with the same data using MLR, variables X_1, X_2 and Y were transformed into standard scores to obtain common units of measurement. Using the following regression models, the hypothesis $H_o: a_1 = a_2$ (where a_1 and a_2 are partial regression weights) was tested.

Full Model 1 VS Restricted Model 2

$z_y = a_1 z_{x1} + a_2 z_{x2} + E_1$ VS $z_y = a_2 z_{x3} + E_2$

(In standard score form z_{y1} represents sex, z_{x1} represents GPA, z_{x2} represents reading achievement and z_{x3} represents the predictor score regardless of whether the person came from sample 1 (s_1) or sample 2 (s_2); $z_{x3} = z_{x1} + z_{x2}$; a_3 represents the common slope for a_1 and a_2.)

Full Model 1

Model 1 $z_y = a_1 z_{x1} + a_2 z_{x2} + E_1$

	z_{y1_1}	z_{x1_1}	0	-
	z_{y1_2}	z_{x1_2}	0	-
s_1	.	.	.	.
	.	.	.	.
	$z_{y1_{30}}$	$z_{x2_{30}}$	0	-
	z_{y1_1}	0	z_{x2_1}	-
	z_{y1_2}	0	z_{x2_2}	-
s_2	.	.	.	.
	.	.	.	.
	$z_{y1_{30}}$	0	$z_{x2_{30}}$	-

Restricted Model 2

Restriction: $a_1 = a_2$

Model 2 $z_y = a_3 z_{x3} + E_2$

	z_{y1_1}	z_{x1_1}
s_1	z_{y1_2}	z_{x1_2}
	.	.
	.	.
	$z_{y1_{30}}$	$z_{x1_{30}}$
	z_{y1}	z_{x2_1}
	z_{y1_2}	z_{x2_2}
s_2	.	.
	.	.
	$z_{y1_{30}}$	$z_{x2_{30}}$

Results for Study One

Testing Model 1 against Model 2 will determine if $a_1 \neq a_2$. The testing of Model 1 against Model 2 should give the same results as one would get by using formula one, the z test.

Reported in Table 1 are F values obtained by testing Model 1 against Model 2 for the 100 pairs of samples drawn (F critical for $df_1 = 1$, $df_2 = 28$, $\alpha = .05 = 3.34$).

Only four of the F values computed when testing Model 1 against Model 2 exceeded the critical value or for this problem four times in a hundred a null hypothesis was rejected when alpha was set at .05.

When the z and F scores in Table 1 were compared, it was found that in 98 percent of the cases the same conclusion would have been drawn regarding the hypothesis $H_o: r_1 - r_2 = 0$. For two of the cases in which the F scores exceeded the critical value, this was also true of the z scores. Examination of cases 44 and 80 show the F scores exceeded the critical magnitude while the z scores narrowly failed to reach significance. (Critical z = 1.96, observed z scores were 1.88 and 1.86 respectively.)

Table 1

Comparison Data for Independent Samples Testing the Hypothesis that $r_1 - r_2 = 0$ Using MLR Vs Z Test

Sample	R_{12}	R_{13}	F	Z
1	-.0382	-.1445	.2275	.3906
2	.1097	-.0822	1.0266	.7051
3	.1659	.2998	.0562	-.4921
4	-.1066	-.0515	.0213	-.2024
5	.2092	.1476	.0341	.2264
6	-.2958	-.2087	.2870	-.3202
7	-.2195	-.2314	.0277	.0439
8	.1568	.1993	.0074	-.1563
9	-.0877	-.3084	.0018	.8110
10	.2246	-.4637	7.8551	2.5292
11	-.0876	.0219	.2493	-.4023
12	-.3548	-.0749	.8628	-1.0285
13	.0987	.1240	.0018	-.0932
14	-.2053	.0243	.7295	-.8433
15	-.1435	-.1563	.0075	.0473
16	-.0480	.1629	.5986	-.7749
17	.0000	-.1785	.6556	.6557
18	-.2496	.0925	1.8128	-1.2570
19	.1861	-.0689	.8114	.9370
20	-.1435	-.1249	.0067	-.0682

Table 1
(Continued)

Sample	R_{12}	R_{13}	F	Z
21	-.2176	-.3741	.2903	.5749
22	.0697	-.0965	.3642	.6106
23	.1512	-.0695	.7133	.8108
24	-.4703	-.2006	1.6368	-.9910
25	.0205	.1436	.2862	-.4523
26	-.1161	-.3024	.3117	.6844
27	.0474	-.0715	.1613	.4369
28	-.0955	-.2009	.2045	.3874
29	.1499	.0943	.0608	.2045
30	-.2011	.0636	1.0676	-.9725
31	.0724	-.2793	1.9379	1.2923
32	-.2372	-.3728	.1464	.4984
33	-.3727	-.1552	.9003	-.7990
34	-.0693	.1093	.4751	-.6564
35	-.1826	-.1275	.1153	-.2021
36	.2586	-.3983	7.2476	2.4136
37	.0553	.1837	.7674	-.4716
38	-.1205	-.1058	.0764	-.0540
39	-.1014	-.1708	.0305	.2550
40	-.1141	-.0091	.0876	-.3857
41	-.0957	.0036	.2218	-.3650
42	.1683	-.1274	1.4081	1.0867

(Continued)

mple	R_{12}	R_{13}	F	Z
3	.0874	.2461	.7123	-.5830
4	-.2835	.2293	4.6929	-1.8841
5	.0879	-.2377	1.8545	1.1964
6	-.2411	-.0628	.6567	-.6553
7	.0388	.1021	.0292	-.2326
8	-.0666	-.1132	.0020	.1710
9	-.1985	.0159	.6105	-.7879
0	-.1084	.3008	2.1849	-1.5033
1	-.1756	-.2217	.0206	.1692
2	-.2905	-.3968	.3242	.3905
3	-.2900	-.3334	.0202	.1592
4	-.0976	.0574	.3046	-.5696
5	-.3273	.0132	1.6870	-1.2510
6	-.0844	-.3070	.9242	.8181
7	-.1926	-.1973	.2412	.0175
8	.0219	.2391	1.3448	-.7980
9	-.2871	-.2601	.0793	-.0992
0	-.4133	-.2789	.2876	-.4938
1	.2297	.0214	.5132	.7654
2	-.2553	-.0265	.4226	-.8410
3	.0067	-.0732	.0108	.2935

Table 1
(Continued)

Sample	R_{12}	R_{13}	F	Z
64	-.1030	-.1145	.0104	.0424
65	-.1915	-.0339	.6764	-.5789
66	-.2034	-.1349	.0401	-.2517
67	-.1258	-.2431	.0679	.4310
68	-.3065	-.1268	.3624	-.6604
69	-.0125	-.1422	.2748	.4766
70	-.2968	.1366	2.6689	-1.5924
71	-.1137	-.1915	.0614	.2859
72	-.2050	-.1759	.0064	-.1069
73	-.0795	.0732	.2265	-.5611
74	.2440	-.2026	2.7359	1.6410
75	.1383	-.0037	.3197	.5215
76	.0074	-.0249	.0156	.1189
77	.0569	.0548	.0231	.0079
78	-.2532	-.3867	.4696	.4907
79	-.1658	.1182	.8270	-1.0434
80	-.3158	.1903	3.7560	-1.8594
81	-.2261	.0801	1.5966	-1.1251
82	-.2011	-.0749	.2890	-.4638
83	-.0156	.0234	.0150	-.1433
84	-.1663	-.0888	.1434	-.2845

Table 1
(Continued)

mple	R_{12}	R_{13}	F	Z
35	-.1278	.1714	.9152	-1.0993
36	-.1424	-.0176	.0748	-.4583
37	-.3278	.1275	3.1805	-1.6732
38	-.0852	.0105	.1551	-.3515
9	-.1290	-.0545	.1322	-.2736
0	-.0962	-.0521	.0419	-.1620
1	.0666	-.0978	.3745	.6041
2	-.1241	-.3654	.6108	.8869
3	.0000	-.1096	.7583	.4025
4	-.2129	.0429	.7594	-.9398
5	-.3984	.0154	1.8932	-1.5204
6	.0182	.1301	.0856	-.4111
7	.0736	.0088	.0604	.2383
8	-.2708	-.0433	1.1240	-.8360
9	.0142	-.1267	.3515	.5177
0	-.0968	-.1105	.0084	.0506

e: The F and Z values for the 100 samples were in agreement 98% of the time. Data for sample 44 and 80 showed the F with 1 and 28 df to be significant while the Z value narrowly fail to be significant. Critical value of F was 3.34.

Comparison of the Glass and Stanley Procedure (z test) to the Hinkle, Wiersma and Jurs Procedure (t test) to the Newman Procedure (MLR) in Testing the Hypothesis $H_o: r_1 - r_2 = 0$, $H_A\ r_1 - r_2 \neq 0$ $\alpha = .05$ for Dependent Sample Data.

Method for Study Two

Solution to the problem of testing for a significant difference between r_1 and r_2 when dependent samples are used must take into account the lack of orthogonality by including the degree of co-variance between the two samples in the error term of the test. Results of solving this problem using the three procedures referred to will be reported below.

A Monte Carlo procedure was used to draw 100 pairs of dependent samples. Correlations were run between similar predictor (X) and criterion (Y) variables in each sample (r_1 and r_2). The criterion variable (Y) was the dichotomous variable sex. Predictor variables were GPA (X_1) and reading achievement (X_2). All data were obtained from the Coleman Data Bank.

Formula 2, presented below, is the solution suggested by Glass and Stanley, 1970.

Formula 2

$$z = \frac{N(r_{xy} - r_{x2})}{\sqrt{(1-r^2_{xy})^2 + (1 - r^2_{xz})^2 - 2r^3_{yz} - (2r_{yz} - r_{xy}r_{xz} \ldots}}$$

$$(1 - r_{xy} - r^2_{xz} - r^2_{yz})$$

Inspection of z scores obtained using Formula 2 and reported in Table 2 indicate that two of the 100 tests reached the critical value of 1.96. Thus, in only two cases was the null hypothesis rejected.

Formula 3 presented below is the solution suggested by Hinkle, Wiersma and Jurs, 1979.

Formula 3

$$t = \frac{(r_{xy} - r_{x2})\sqrt{(n-3)(1+r_{yz})}}{\sqrt{2(1 - r^2_{xy} - r^2_{xz} - r^2_{yz} + 2r_{xy}r_{xz}r_{yz})}}$$

Inspection of t scores obtained by using Formula 3 and reported in Table 2 indicates that two of the 100 tests reached the critical value of 2.00 with df = 57. Thus, in only two cases was the null hypothesis rejected.

The MLR procedure used to test for a significant difference between r_1 and r_2 obtained from dependent samples involved the transformation of predictor and criterion variables into standard scores in order to obtain common units of measurements. The hypothesis that $H_o: a_1 = a_2 = a_3$ (a_3 represents the common slope) was tested by comparing the amount of variance accounted for by the following regression models. Values a_1, a_2, a_3 are partial regression weights.

Full Model 3 VS Restricted Model 4

Theoretical Models 3 and 4

$$z_y = a_1 z_{x1} + a_2 z_{x2} + a_4 p_1 + \dots + a_{60} p_{63} + E_3$$

VS

$$z_y = a_3 z_{x3} + a_4 p_1 + \dots a_{63} p_{60} + E_4$$

Peddhazur's conceptual approach for Models 3 and 4 where small ps are collapsed and designated as a large P. (See Peddhazur, 1977; Williams, 1977.)

$$z_y = a_1 z_{x1} + a_2 z_{x2} + a_4 P + E_3 \text{ VS } z_y = a_3 z_{x3} + a_4 P + E_4$$

(In standard score form z_{x1} represents GPA in sample 1 (s_1) z_{x2} represents reading achievement for the same persons in sample 2 (s_2), z_{x3} represents the predictor score regardless if the score came from sample 1 or 2; $z_{x3} = z_{x1} + z_{x2}$ and z_y represents the criterion variable sex; a_4 is a partial regression weight; Ps represent person vectors used to account for the co-variance between the two dependent samples; a_3 represents the common slope for the partial regression weights a_1 and a_2.)

Below is a simulated numerical example to explain the procedure.

Full Model 3

Model 3		z_y =	$a_1 z_{x1}$ +	$a_2 z_{x2}$ + a_4	P + E_3
Sub.	1	1	1	0	2.5
	2	1	.5	0	1.2
s_1	3	0	-.3	0	-.5
	4	0	.7	0	1.6

Sub.	1	0	0	1.5	2.5
	2	0	0	.7	1.2
s_2	3	1	0	-.2	-.5
	4	1	0	.9	1.6

Restricted Model 4

Restriction $a_1 = a_2 = a_3$

Model 4 $z_y = a_3 z_{x3} + a_4 P + E_4$

Sub.	1	1	1	2.5
	2	1	.5	1.2
s_1	3	0	-.3	-.5
	4	0	.7	1.6
Sub.	1	0	1.5	2.5
	2	0	.7	1.2
s_2	3	1	-.2	-.5
	4	1	.9	1.6

Attention is directed to the procedure used to develop the person vectors. Model 3 represents the prediction of sex (z_y) by the standardized GPA (z_{x1}), the standardized reading achievement score (z_{x2}) and a composite person vector (P). In the simulated model there are four subjects, two males and two females, each of whom is measured twice; once on GPA and once on reading achievement. The person vector is then computed by adding the GPA score of subject 1 to the reading achievement score of subject 1; which in the simulated case sums to 2.5. Similarly for subject 2 one adds GPA to reading achievement and places the total 1.2 in the two positions of the person vector representing subject 2. This procedure is repeated until all persons are represented by a person vector, thus accounting for the covariance between the dependent samples.

Results for Study Two

Reported in Table 2 are F values obtained by testing Model 3 against Model 4 for the 100 samples drawn (F critical for $df_1 = 2$, $df_2 = 57$, $\alpha \doteq .05$ is 3.17). Only two of the F values computed exceeded the critical value. Thus, for only two cases was the null hypothesis rejected with alpha set at .05.

When the z, t and F scores reported in Table 2 were compared, it was found that for the same two cases (58 and 62) the z, t and F test results were significant. It is, therefore, apparent that there was 100 percent agreement among the three procedures used.

Conclusion

To the extent that the approaches suggested by Minium, 1978; Glass and Stanley, 1970; Hinkle, Wiersma and Jurs, 1979 are valid, the use of multiple linear regression has been demonstrated to be a viable procedure for testing for a significant difference between r_1 and r_2 with both dependent and independent data. Results using MLR were in

Table 2

Comparison Data for Dependent Samples Testing the Hypothesis that $r_1 - r_2 = 0$ Using MLR (F) Vs the Z Test Vs the t Test

Sample	R_{12}	R_{13}	R_{23}	F	Z	t
1	-.1081	-.1019	-.5947	.0141	-.0273	-.0271
2	-.0049	-.0730	-.4437	.0600	.3109	.3033
3	-.0685	-.1147	-.4848	.0200	.2094	.2059
4	-.2429	.1196	-.4731	2.0612	-1.7047	-1.6438
5	-.0427	-.1308	-.3334	.1386	.4213	.4125
6	-.2337	-.0451	-.5143	.5423	-.8629	-.8587
7	-.1578	-.1723	-.4318	.0002	.0680	.0680
8	-.1443	-.0763	-.4686	.1262	-.3113	-.3071
9	-.1229	-.2118	-.4150	.1868	.4203	.4201
10	-.0698	-.1742	-.5636	.0899	.4648	.4625
11	.0388	-.2175	-.5485	.9222	1.1573	1.1320
12	-.2294	.1421	-.5152	2.1572	-1.7221	-1.6558
13	-.0391	-.0373	-.5848	.0003	-.0080	-.0078
14	-.2290	-.0820	-.0279	.1535	-.8143	-.7983
15	-.0909	-.0583	-.5034	.0285	-.1466	-.1438
16	-.0975	-.1126	-.3899	.0074	.0709	.0697
17	-.1112	-.0085	-.4717	.2554	-.4666	-.4559
18	-.0344	.0036	-.4687	.0137	-.1719	-.1675
19	.0208	-.0351	-.6242	.0499	.2406	.2344

Table 2

(Continued)

Sample	R_{12}	R_{13}	R_{23}	F	Z	t
20	-.1637	.0648	-.6433	.6383	-.9927	-.9661
21	-.1286	-.0124	-.5309	.1273	-.5187	-.5079
22	.1160	-.1413	-.3697	.9576	1.2280	1.1886
23	-.1486	-.2208	-.3999	.1420	.3449	.3464
24	-.0933	.0405	-.5993	.2844	-.5828	-.5674
25	-.1714	.0376	-.5277	.5597	-.9417	-.9184
26	.1634	-.1349	-.5247	1.2434	1.3573	1.3096
27	-.1369	.0733	-.2734	.6172	-1.0346	-1.0046
28	.0417	-.1020	-.4548	.2704	.6571	.6395
29	.0000	-.0934	-.3753	.0726	.4381	.4273
30	.0378	.2072	-.3869	.4074	-.8049	-.7919
31	-.1102	.2080	-.4317	1.5068	-1.5030	-1.4518
32	-.1122	.0247	-.3253	.3005	-.6559	-.6389
33	-.1519	.0089	-.5131	.2809	-.7245	-.7084
34	-.1304	.1154	-.5743	.8748	-1.0910	-1.0559
35	.1613	-.2518	-.5863	2.4361	1.8883	1.8093
36	-.2672	.1156	-.4620	2.2825	-1.8183	-1.7541
37	-.1817	.1420	-.6488	1.5265	-1.4215	-1.3691
38	-.2039	-.2760	-.3791	.3468	.3548	.3639
39	-.1299	-.1480	-.3844	.0000	.0858	.0849
40	.0268	-.2342	-.4445	1.0715	1.2244	1.1973

(Continued)

Sample	R_{12}	R_{13}	R_{23}	F	Z	t
41	-.0862	.0675	-.5479	.3316	.6814	-.6624
42	-.1134	-.1388	-.3589	.0068	.1214	.1197
43	-.1972	-.0269	-.5492	.4731	.7643	.7557
44	.0722	-.1105	-.4592	.5056	.8365	.8126
45	.1039	-.0568	-.4406	.3424	.7391	.7186
46	-.1462	.0208	-.5723	.3459	-.7379	-.7210
47	-.1396	.0224	-.4709	.3753	-.7395	-.7212
48	-.1728	-.1842	-.6266	.0000	.0502	.0521
49	-.1147	-.0868	-.5359	.0033	-.1245	-.1230
50	.0794	.0287	-.5166	.0303	.2263	.2212
51	-.0849	-.0789	-.3848	.0007	-.0280	-.0274
52	-.1164	-.1037	-.4168	.0085	-.0591	-.0582
53	-.1752	-.0219	-.5276	.2925	-.6898	-.6789
54	-.2124	-.0444	-.4226	.4027	-.7889	-.7785
55	.0896	-.0502	-.3565	.2555	.6618	.6439
56	.1172	-.2122	-.5267	1.5152	1.5087	1.4563
57	-.1864	.1915	-.6415	2.0257	-1.6806	-1.6098
58	-.4106	.0709	-.4124	4.1282	-2.4361	-2.3888
59	-.0425	-.0898	-.3059	.0451	.2278	.2226
60	.0822	-.1186	-.5032	.6037	.9077	.8809
61	-.2270	.0985	-.3751	1.5038	-1.5736	-1.5216
62	-.1543	.3160	-.6016	3.6028	-2.1797	-2.0931

Table 2

(Continued)

Sample	R_{12}	R_{13}	R_{23}	F	Z	t
63	.0251	-.1084	-.4283	.2451	.6155	.5996
64	-.2392	-.0805	-.3373	.3583	-.7738	-.7665
65	-.2249	-.2041	-.4498	.0027	-.0986	-.1010
66	-.1786	-.0302	-.2974	.4226	-.7250	-.7099
67	-.1096	.1093	-.3481	.6796	-1.0474	-1.0157
68	-.0522	.0362	-.4824	.1157	-.3990	-.3885
69	-.1247	.1739	-.3957	1.3316	-1.4216	-1.3729
70	.0057	-.1372	-.6823	.3335	.6094	.5983
71	-.1623	.1125	-.3261	1.0284	-1.3374	-1.2939
72	-.1065	-.1328	-.4727	.0680	.1202	.1190
73	.0209	.0120	-.4677	.0018	.0405	.0395
74	-.2765	-.0479	-.3575	.5527	-1.1157	-1.1045
75	.1378	-.0768	-.3822	.7065	1.0143	.9844
76	-.0664	-.1420	-.3096	.1457	.3661	.3591
77	-.1462	-.1286	-.3292	.0466	-.0850	-.0840
78	-.2009	-.0959	-.5044	.3175	-.4800	-.4800
79	-.0960	-.2412	-.2796	.1335	.7244	.7171
80	-.1449	.0812	-.4936	.7979	-1.0293	-.9985
81	-.1217	-.1347	-.4893	.0955	.0591	.0587
82	-.0315	-.2108	-.5069	.3846	.8179	.8087
83	-.1265	.0388	-.4502	.4362	-.7589	-.7388
84	-.1869	.1276	-.5317	1.4378	-1.4325	-1.3818

Table 2

(Continued)

Sample	R_{12}	R_{13}	R_{23}	F	Z	t
85	-.1404	.0383	-.6209	.5040	-.7773	-.7583
86	-.1376	-.0578	-.4992	.1262	-.3607	-.3552
87	-.2128	.1423	-.4939	1.7676	-1.6518	-1.5890
88	-.1787	-.0716	-.5486	.0233	-.4796	-.4772
89	-.0400	-.0773	-.6466	.0179	.1600	.1570
90	-.0700	.0425	-.5035	.1760	-.5043	-.4909
91	-.0341	-.2049	-.5194	.3524	.7750	.7667
92	-.0233	-.2465	-.4014	1.0884	1.0644	1.0487
93	-.0624	-.0942	-.3143	.0317	.1524	.1491
94	.1593	-.0670	-.5301	.7893	1.0181	.9892
95	-.1582	-.1596	-.3849	.0001	.0067	.0067
96	.0196	-.3220	-.6075	2.6945	1.5569	1.5625
97	-.2074	.0764	-.4796	1.1570	-1.3129	-1.2740
98	-.0962	-.0658	-.5215	.0722	-.1357	-.1334
99	-.1564	.0108	-.5323	.3164	-.7495	-.7332
100	-.2481	.0265	-.4969	.9063	-1.2694	-1.2449

Note: The F, Z and t values for the 100 samples were in agreement 100% of the time. Degrees of freedom for the F and t values were $F_1 = 2$, $F_2 = 57$ and $df_t = 57$. The critical value of F was 3.17 for t it was 2.00.

98 percent agreement with the procedure suggested by Minium (1978) for dependent data. For the two cases (44 and 80) where the MLR results did not agree with the more traditional procedure, the observed values just missed reaching the critical level, 1.88 and 1.86 respectively. When data from dependent samples were evaluated, there was 100 percent agreement among the procedures suggested by Glass and Stanley, 1970 (z test); Hinkle, Wiersma and Jurs, 1979 (t-test); and Newman (MLR).

The similarity in the results tends to support the use of all procedures tested. The writers, however, found the traditional tests (z and t) to be more cumbersome when a computer program for testing general linear models was available. In addition to the pragmatic consideration, a pedagogical advantage seems to exist when using MLR. Teaching students how to use the general linear model permits them to conceptualize more clearly what they are doing. This would be especially true for more naive students for whom application of the traditional models may be based entirely upon what appears to be unrealted statistical procedures. For the more sophisticated individual, MLR facilitates expression of the research question of interest in terms of general linear models without having to worry about a specific procedure to use for that particular problem.

Further, it is the belief of the authors that the general linear model approach to testing hypotheses is more apt to increase the ability of the researchers to ask questions that are of most specific interest to them; reducing the likelihood of their making a Type VI Error, Newman, I.; Deitchman, R.; Burkholder, J.; Sanders, P.; and Ervin, L. (1976) and Roll, S.; Hoedt, K.; and Newman, I. (1979). A Type VI Error is the inconsistency between the research question of interest and the statistical model being applied.

References

Dayton, M. C. & Stunkard, C. L. Statistics for problem solving. New York: McGraw-Hill, Inc., 1971.

Glass, G. V. & Stanley, J. C. Statistical methods in education and psychology. Prentice-Hall, Inc., 1970.

Hinkle, D. E.; Wiersma, W. & Jurs, S. G. Applied statistics for the behavioral sciences. Rand McNally College Publishing Company, 1979.

McCall, R. B. Fundamental statistics for psychology. New York: Harcourt, Brace & World Books, Inc., 1970.

Minium, E. W. Statistical reasonsing in psychology and education. New York: John Wiley & Sons, Inc., 1978.

Newman, I.; Deitchman, R.; Burkholder, J.; Sander, R.; & Ervin, L. Type VI Error: Inconsistency between the statistical procedure and the research question. Multiple Linear Regression Viewpoints, 1976, 6, 1-19.

Peddhauzur, E. J. Coding subjects in repeated measure design. Psychological Bulletin, 1977, 84, 298-305.

Tate, M. T. Statistics in education and psychology. New York: The Macmillan Company, 1965.

Roll, S.; Hoedt, K. C.; & Newman, I. A demonstration of Type VI error: An applied research problem. Multiple Linear Regression Viewpoints, 1979, 10(1), 31-38.

Williams, J. D. A note on coding the subjects' effects in treatments x subjects designs. _Multiple Linear Regression Viewpoints_, 1977, _8_(1), 32-35.

APPLICATION OF JUDGMENT ANALYSIS TO INTERRATER COMPARISONS

Linda E. Kapuniai

University of Hawaii

Joe H. Ward, Jr.

University of Texas at San Antonio

David H. Crowell

University of Hawaii

Michael J. Light and Rodney B. Boychuk

University of Hawaii

Joan E. Hodgman

USC Medical Center

ABSTRACT

A multiple regression method is presented for comparing the bases of two raters' judgments. This technique, which has been referred to as judgment analysis or policy capturing, is described for judgments of two nurses. In the example presented, judgments of future infant performance were derived from the nurse's scoring of infants' behavior on the Brazelton Neonatal Behavioral Assessment Scale. Brazelton dimension scores served as predictors of future performance in a test of differences between the policies (criteria) of the two nurse-raters. Sample data illustrate the technique but do not constitute a direct test of the data since the two nurse's ratings were actually on two different sets of infants. If the ratings had been on the same babies or identical samples of babies, the technique would have revealed, first, that the two nurses based their judgments primarily on one Brazelton dimension, interactive processes; and second, that one nurse consistently rated the babies' future performance at a higher level than did the other nurse. This technique has potential application to evaluation of rating criteria for training of observers or judges and in other problem solving areas such as conflict resolution.

Subjective predictions of progress and objective assessments of behavior are frequently required in many programs and projects. Consistency and accuracy of these observations are important issues in evaluating the judgments of different individuals or policies in relation to patterns of attributes (Fisher, 1983; Most & Starr, 1983). When assessing these judgments several questions often arise, for example, which of the many observations contributed the most to the overall judgment? Or, more importantly, if more than one observer is involved, to what extent did the raters rely on the same criteria as the basis for their predictions.

This paper presents a general statistical method for comparing the observations and determining the bases of the judgments of two individuals. The method is applied to observations on the current status and judgments of future capabilities of newborn infants. The observations were made by two nurses in the process of conducting the Brazelton Neonatal Behavioral Assessment Scale, BNBAS, (Brazelton, 1973). Judgment of the infant's future performance was made after completion of the BNBAS assessment. To illustrate the method, the nurse's judgments are treated as if they were rating the same infants. The two raters' judgments are then compared in terms of the regression weights associated with BNBAS dimension scores (Als, Tronick, Lester & Brazelton, 1977) derived from the original BNBAS observations. These scores represent the following dimensions: 1. Interactive processes: capacity to respond to social stimuli through orientation, cuddling and consoling; 2. Motoric processes: ability to maintain good tone, control motor behavior and integrate actions; 3. Organizational processes: ability to modulate states of consciousness in interactions with the environment primarily by shutting out aversive stimuli; and 4. Physiological reaction to stress: stability in response to stress. Dimensions 1-3 are scored as follows: 1 for good, 2 for average, and 3 for worrisome or deficient performance. Dimension 4 is coded either 1 (good) or 0 (bad).

The general model used as a basis for analysis of degree of rater similarity is a multiple linear regression using least-squares estimation of the regression weights,

$$Y = w_1 U + w_2 X_1 + \ldots + w_3 X_2 + \ldots + w_k X_k + E, \qquad (1)$$

where w_1, w_2, and $w_3 \ldots w_k$ are least squares weights that minimize the squared errors in E. U is a vector of "1"s and X_1 and $X_2 \ldots X_k$ are the K predictor vectors. The dependent variable, Y, is the set of judgments or ratings of the situations characterized by the predictor data. The regression approach outlined here is a variation of a technique called "policy capturing," (Christal, 1968a, b; Christal & Bottenberg, 1969; Ward, 1979). The combination of the regression weights applied to each variable is taken as defining the rater's "policy" with regard to Y, the dependent variable.

The general hypotheses to be tested are: "Does the policy used by one rater differ from that used by another?" and "If the two policies differ, do they differ by a constant amount?"

The models for interrater comparison are presented first followed by their application to sample BNBAS data.

METHOD

Model Development

The following regression equations were designed to test the judgments of the two raters, Nurse 1 and Nurse 2, on the four BNBAS dimension scores. Each nurse's equation would take the general form,

$$Y_{nurse} = \text{function of (Dimension 1, Dimension 2, Dimension 3, and Dimension 4)} + E \qquad (2)$$

where Y is a nurse's judgment of an infant's performance. A similar regression equation is established for Nurse 2.

Model 1. Model 1, which incorporates both nurses' equations into a single model, takes into account the possibility that Nurse 1 makes ratings of infants that yield an equation (weights a_0, a_1, a_2, a_3, a_4) that differs from the corresponding equation (weights b_0, b_1, b_2, b_3, b_4) of Nurse 2. The equation is:

$$Y = a_0P1 + a_1(P1*D1) + a_2(P1*D2) + a_3(P1*D3) + a_4(P1*D4) + b_0P2 + b_1(P2*D1) + b_2(P2*D2) + b_3(P2*D3) + b_4(P2*D4) + E1 \qquad (3)$$

where Y is the vector of future infant performance ratings from both nurses, D1 to D4 are the four BNBAS dimension scores, P1 is "1" for Nurse 1 and 0 otherwise, P2 is "1" for Nurse 2 and 0 otherwise, and E1 is the error in Model 1. In other words, the nurses are assumed to have based their predictions on two completely different policies. The least squares solution for Equation 3 will yield two sets of weights that might be different. Dimension 1 for Nurse 1 (P1*D1) has one weight (a_1) assigned to it, dimension 1 for Nurse 2 (P2*D1) may have another weight (b_1) assigned to it, and so on. Furthermore P1 is assigned one weight (a_0) and P2 may have another weight (b_0).

Model 2. To test the hypothesis that the two nurses' predictions differed by a constant, restrictions are imposed on Model 1 to obtain Model 2, Equation 4. To illustrate this point, we would act as if the hypothesis is: when two nurses are presented with 10 babies and asked to make predictions independently on those 10 babies, the predictions will differ by a constant amount. The restrictions implied by the hypotheses of constant differences are:

$$a_1 = b_1 = c_1,\ a_2 = b_2 = c_2,\ a_3 = b_3 = c_3, \text{ and } a_4 = b_4 = c_4$$

Substituting these restrictions in Model 1 gives Model 2.

$$Y = a_0 P1 + b_0 P2 + c_1 D1 + c_2 D2 + c_3 D3 + c_4 D4 + E2. \qquad (4)$$

Observe that this model has the same weights (c_1, c_2, c_3, c_4) for the two nurses, but that the nurses' judgments will differ by the constant value $a_0 - b_0$.

Model 3. Model 3 assumes that the policies used by Nurses 1 and 2 are identical. The restriction on Model 2 implied by this hypothesis is $a_0 = b_0 = c_0$. Substituting this restriction in Model 2 gives Model 3,

$$Y = c_0 U + c_1 D1 + c_2 D2 + c_3 D3 + c_4 D4 + E3, \qquad (5)$$

where U = P1 + P2, the Unit Vector containing a "1" in every element. Observe that this model has given up all information that distinguishes the two nurses.

Testing the Hypotheses.

After Models 1, 2 and 3 (equations 3, 4, and 5) have been developed, the questions of policy differences can be answered by comparing the R^2's (squared multiple correlations) from the equations. The question, "If the two policies differ, do they differ by a constant amount?" can be answered by determining if R_1^2 is significantly larger than R_2^2. This comparison, Equation 6, is made by calculating

$$F = \frac{(R_1^2 - R_2^2) / (n_1 - n_2)}{(1 - R_1^2) / (N - n_1)} \qquad (6)$$

which is distributed as F with degrees of freedom $(df_1) = (n_1$

$- n_2$) and (df_2) = ($N - n_1$): n_1 (=10) is the number of coefficients in Model 1, n_2 (=6) is the number of coefficients in Model 2 and N (=45) is the total number of ratings by both nurses. If the F-test is not significant we accept the restricted Model 2, that is the hypothesis that the differences between the two nurse's policies are constant is not rejected. The difference will be (a_0 - b_0). In this case Model 2 would be adopted.

The next step in the analysis depends on the result of the comparison between Model 1 and Model 2. If we reject the constant difference hypothesis we conclude that the policies differ, and, therefore, Model 1 is appropriate.

If we accept the constant difference hypothesis Model 2 is assumed, and to test that the policies are identical we compare Model 2 with Model 3 as in equation (7),

$$F = \frac{(R_2^2 - R_3^2) \ / \ (n_2 - n_3)}{(1 - R_2^2) \ / \ (N - n_2)} \qquad (7)$$

where R_2^2 is compared to R_3^2. If R_2^2 is significantly larger than R^2, the null hypothesis ($a_0 = b_0 = c_0$) is rejected and it can be concluded that the nurses differ in their ratings and the difference is constant. If the difference in the two R^2 is not significant, it is concluded that there are no differences between the nurses' judgments when expressed in terms of the four BNDAS dimensions.

Model Application.

Subjects and Procedure. Subjects were 45 infants who were seen at term as part of a larger study of metabolic derangements, neurophysiological functioning and behavior. Informed consent was obtained from parents and physicians prior to testing. Brazelton assessments for 25 of these infants were conducted by one rater, Nurse 1, and the remaining 20 by a second rater, Nurse 2. The same assistant recorded the scores during the BNBAS tests done by both nurses. After each test was completed, the test information was combined to form the four dimension scores (Als et al., 1977). Subsequent to the determination of the four dimension scores the nurses made a Judged Future Performance, JFP, for each infant. This JFP was scored as 0, 1, or 2, to correspond with predictions of below average, average, or above average future performance. No other explicit criteria were suggested.

Results The scores for the four dimensions resulting from the test of the two nurses are in Table 1.

Table 1. Judged Future Performance (JFP), and Brazelton Neonatal Assessment Scale Dimension Scores From Two Nurses

Nurse 1						Nurse 2					
Case	JFP	D1	D2	D3	D4	Case	JFP	D1	D2	D3	D4
1	2	1	2	1	1	26	2	1	2	1	1
2	1	2	1	2	1	27	1	2	1	2	1
3	1	1	2	2	1	28	1	2	2	2	1
4	1	2	2	2	1	29	1	3	2	2	1
5	1	2	2	3	1	30	2	1	2	2	1
6	1	1	1	2	1	31	1	2	2	3	0
7	2	1	2	1	1	32	1	3	2	2	1
8	1	2	2	2	1	33	1	3	2	2	1
9	1	2	2	3	1	34	1	3	2	2	1
10	1	2	1	2	1	35	1	2	2	2	1
11	1	1	2	2	0	36	2	1	1	2	1
12	1	2	1	2	1	37	1	2	2	2	1
13	0	3	1	2	0	38	2	2	1	2	1
14	1	1	3	3	1	39	1	3	2	2	1
15	1	3	2	2	0	40	1	3	2	2	1
16	1	3	2	2	1	41	1	2	2	2	1
17	1	3	1	2	1	42	1	1	2	2	1
18	1	2	1	2	1	43	2	1	1	2	1
19	0	3	3	2	1	44	2	2	2	2	1
20	2	1	2	1	1	45	2	2	2	2	1
21	1	2	2	1	1						
22	1	3	2	3	1						
23	1	3	2	2	1						
24	2	1	2	1	1						
25	1	1	1	1	1						

The four dimension scores, nurse identification and ratings of future performance were then entered into the models previously described. The results were $R_1^2 = 0.931$; $R_2^2 = 0.926$; and $R_3^2 = 0.912$. The R^2 values were entered into the $\underline{F}$-test formulas with the appropriate degrees of freedom. First, Model 1 was compared with Model 2 using Equation 6.

$$\text{Test 1: } \underline{F}_{(4,35)} = \frac{(0.931-0.926) \;/\; (10-6)}{(1 - 0.931) \;/\; (45-10)} = .597 \qquad (8)$$

Test 1 (Model 1 compared with Model 2) was not significant. In light of this result, Model 2 was assumed where $a_1 = b_1 = c_1$, $a_2 = b_2 = c_2$, $a_3 = b_3 = c_3$ and $a_4 = b_4 = c_4$. Since Test 1 indicated that nurses' judgments differed by a constant amount, Model 2 was compared to Model 3 in Test 2, equation (9), using equation (7) above.

$$\text{Test 2: } \underline{F}_{(1,39)} = \frac{(0.926-0.912) \;/\; (6-5)}{(1 - 0.926) \;/\; (45-6)} = 7.26 \qquad (9)$$

The $\underline{F}$ of 7.26 was significant at $\underline{p} < .05$; therefore, the null hypothesis, that $a_0 = b_0 = c_0$, was rejected. While the expected nurses' ratings of future performance differed by a constant amount, the constant difference was not zero. The estimate of the actual difference was $a_0 - b_0 = 1.93 - 2.24 = -.31$ (see Table 2).

Table 2. Regression Results for Model 2 (Equation 4)

Predictor	Coefficient	F-Value* (df=1,39)	Probability
P1-Nurse One	1.93		
P2-Nurse Two	2.24		
D1-Interactive Processes	-.32	16.75	.0002
D2-Motoric Processes	-.02	.04	.8471
D3-Organizational Processes	-.22	3.45	.0708
D4-Physiological Reaction to Stress	.26	1.63	.2098

*F-Values result from the (1,39 degrees of freedom) test that the corresponding coefficient is equal to zero

Since the differences between ratings were constant (Test 1), we can conclude that the relationships between the four BNBAS scores and the judgments of Nurse 1 did not differ from the relationships of Nurse 2. But Test 2 indicated that even though differences were constant there was a significant difference between the level of ratings of the two nurses. Nurse 2 tended to give higher ratings (.31) than Nurse 1.

Since the nurses did not actually rate the same infants it cannot be determined whether these results reflect actual differences in the nurse's policies or differences in the two sets of infants. In this example the relationship of the four BNBAS scores for the judgments was the same for the two nurses; therefore, it was of interest to examine each of the four coefficients c_1, c_2, c_3, c_4. Inspection of the Model 2 regression equation in Table 2 reveals that the two nurses based their judgments primarily on dimension 1 (Interactive Processes). This conclusion is based on the small probability (p = .0002) associated with the hypothesis that babies who have the same scores on Dimension 2, 3, and 4, but different Dimension 1 ratings will have the same expected JRP ratings. The probability of .07 associated with the test on Dimension 3 indicates that Organizational Processes also may contribute to the judgment process.

DISCUSSION

Evaluation of judgments of behavior based on observations is a situation that occurs frequently. It is important not only to know on what bases and how consistently the observer is making judgments, but also whether judgments of different observers or raters have similar bases. Techniques which address these questions are demonstrated in Test 1 and Test 2, multiple regression models which have been described as policy capturing. This approach describes the set of variables or observations that best characterize a judgment.

One possible application of judgment analysis or policy capturing would be training programs where the goals are to evaluate and increase degree of intra- and inter-rater reliability. If the policy or combination of independent variables (observations), does not account for a significant proportion of the variance in the dependent variable, it can be inferred that the judgment of the observer is, to a large degree based on information other than that contained in the predetermined set of observations. In other words, the person is utilizing information not summarized in the behaviors represented by the values of the independent variables in the equation. For example, if the observer is instructed to make an assessment of an infant's future performance based on the results of the BNPAS, and the BNBAS values do not support or predict the JFP, it may be that knowledge of the child's home environment or some other unknown factor was entering into this judgment. In this situation, it may be necessary to retrain the observer to eliminate other than specified information or it may be more desirable to reconsider the factors in the equation. If two raters (judges or observers) differ in their rating criteria, the criteria of the rater whose judgments best approximate actual future performance can be adopted as the standard for others. These same considerations could be pertinent to questions of conflict resolution, both in refining the dependent variable (Most & Starr, 1983) and as a way of describing how decisions are arrived at in problem-solving or negotiation settings (Fisher, 1983).

CONCLUSION

This technique can be a valuable aid for detecting implicit weightings of unknown variables which result in unexplained variance in judgments, and for standardizing judgments, that is, insuring that they are based on the same criteria.

REFERENCES

Als, H., Tronick, E., Lester, B.M., and Brazelton, T.B. The Brazelton Neonatal Behavioral Assessment Scale (BNBAS), *Journal of Abnormal Child Psychology*, 1977, *5*, 215-230.

Brazelton, T.B. *Neonatal Behavioral Assessment Scale*, Clinics in Developmental Medicine, No. 50, London: S.I.M.P. with Heinemann Medical, Philadelphia. Lippincott, 1973.

Christal, R.E. JAN: A technique for analyzing group judgment, *The Journal of Experimental Education*, 1968a, *36*, 24-27.

Christal, R.E. Selecting a Harem-and other applications of the policy capturing model, *The Journal of Experimental Education*, 1968b, *36*, 35-41.

Christal, R.E., and Bottenberg, R.A. Grouping criteria - A method which retains maximum predictive efficiency, *The Journal of Experimental Education*, 1968, *36*, 28-34.

Fisher, R.J. Third party consultation as a method of intergroup conflict resolution, *Journal of Conflict Resolution*, 1983, *27*, 301-334.

Most, B.A., and Starr, H. Conceptualizing "war" consequences for theory and research, *Journal of Conflict Resolution*, 1983, *27*, 137-159.

Ward, J.H., Jr. Creating mathematical models of judgment processes: from policy capturing to policy specifying, The Journal of Experimental Education, 1979, 48, 60-84.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Volume 13, Number 1 Spring 1984

Judgment Analysis for Characterizing Campus Ministry Function

Barbara Whittington

Southern Baptist Convention

Rex Leonard and W. Lee Pierce

University of Southern Mississippi

Abstract

The purpose of the study was to investigate the judgmental policies of campus ministry held by campus ministers at state-supported universities when the campus ministers were grouped according to the campus minister's ministry group, years of personal campus ministry experience, size of the student body, campus minister's position at the school, and the campus minister's age by decade of birth. The ultimate goal of the research was to provide both clergy and laity with a clearer understanding of the role of campus ministry at state-supported universities. The questionnaire used in the study was developed using the critical incident technique. Supervisors of campus ministry were asked to list the three most important ministry goals or role functions or campus ministry at state-supported universities. The responses were tabulated and a 17-item questionnaire was formed. In order to determine reliability, a pilot test of the questionnaire was conducted. The subjects (N = 276) who participated in the study be responding to the questionnaire were campus ministers in ten ministry group affiliations at state-supported universities during 1982. They rated 17 goals of campus ministry and gave a rating to a program of campus ministry that would have the 17 goals as principal objectives. The Judgment Analysis technique was used and the campus ministers were found to be clustered in six judgmental areas related to ministry group.

The 1969 Wesley Foundation study found that clarification of ministerial roles and the search for self-image were among the greatest concerns of campus ministers (Underwood, 1969). Campus ministry has been in existence long enough to have a very large professional staff and a physical presence on hundreds of campuses (Johnson, 1979). Although this specialized ministry has produced several generations of practitioners and many generations of clients, it is still unable to define its role (Hammond, 1979).

Lanagan (1979) suggested that both the university and the church are involved in determining the role campus ministry plans on campus. The university sees campus ministry as an academic or student life force and asks what preparation the campus minister should have to serve and assist the college or university in achieving its goals. The religious organization with which the campus minister is affiliated sees the campus ministry as a component which fosters a religious atmosphere in the University.

The purpose of this research was to develop purpose statements that could be identified by campus ministers as being relevant to campus ministry and analyze the purpose statements according to the campus minister's ministry affiliation, size of student body, and campus minister's age. The sets of purpose statements can be utilized to provide both clergymen and laity with a clearer understanding of the role of campus ministry at state-supported universities and to provide educational organizations affiliated with campus ministry with direction in planning continuing educational opportunities for campus ministers.

Procedures

The critical incident technique (Flanagan, 1954) was used to develop the instrument used in this study. One-hundred seventy-one supervisors of campus ministry were asked to state what they considered to be the three most important goals of a viable campus ministry. The responses of the supervisors were tabulated and the most frequent responses were used as items (goals) on the insturment (see Table 1).

Table 1

Goals of Campus Ministry

Number	Statement of Goal	Short Title
1.**	To assist students in developing Biblically based life goals and in the integration of these into the vocation of their choice.	Biblically based life goals.
2.**	To provide opportunities for fellowship.	Fellowship.
3.**	To provide worship opportunities on campus.	Worship.
4.*	To develop student leadership.	Student leadership.
5.***	To lead students and faculty to become involved in the local church.	Involved in the local church.
6.***	To nurture students who are considering the religious profession as a vocation.	Religious vocation.
7.*	To expand the vision of students to invest their lives in meeting the needs of a hurting world.	Invest in hurting world.

Number	Statement of Goal	Short Title
8.*	To organize groups for study and action upon special concerns and problems raised in the university.	Organize for study and action.
9.**	To assist persons in their search for religious identity.	Religious identity.
10.**	To provide opportunities for study in doctrine, religious beliefs, and church (denominational) policy.	Study of religious topics.
11.*	To provide students with opportunities for personal ministry.	Personal minstry.
12.*	To nurture students and faculty in faith development.	Faith development.
13.*	To create an environment (organizational structure) in which students can grow in their faith.	Environment for growth.
14.*	To develop a visible community of faith on campus.	Visible community of faith.
15.**	To provide pastoral counseling.	Pastoral counseling.
16.*	To help students and faculty relate their work in academia and in the larger world beyond the campus.	Relate faith.
17.*	To enable the faith community on campus to be able to share their faith with others on campus while respecting the beliefs, values, and lifestyles of those other people.	Sharing of faith.
18.	Assuming that all the foregoing are principal objectives for a campus ministry program, how valid would you judge the overall goal of that ministry to be?	Overall rating of goals.

* Factor 1: Developmental Role of Campus Ministry
** Factor 2: Supportive Role of Campus Ministry
*** Factor 3: Denominational Identity Role of Campus Ministry

Each item on the instrument was scored from one to five. An item received a score of 5 if the dimension was scored as being very important to campus ministry; 4 if the dimension was scored as being of more than average importance; 3 if the dimension was scored as being of average importance; 2 if the dimension was scored as being of less than average importance; and 1 if the dimension was scored as being of little or no importance.

Construct validity of the instrument was investigated using factor analysis. Three factors (constructs) were found to exist and are indicated in Table 1. They were Developmental Role of Campus Ministry, Supportive Role of Campus Ministry, and Denominational Identity Role of Campus Ministry.

The instruments were then mailed to 500 randomly selected campus ministers serving at state-supported universities. The participants were selected from 3,427 campus ministers whose names appeared on mailing lists obtained from the headquarters of National Campus Ministry groups. There were 276 usable responses and Table 2 shows the ton groupings by ministry affiliation.

The sample consisted of 226 males and 50 females and were distributed among four age categories (see Table 2). Almost 64% of the campus ministers were less than 43 years of age. The samplo was further categorized by the size of the student body at the institution where the campus ministry was located (see Table 2). Over 65% of the campus ministries were located at campuses having more than 9,000 students.

Table 2

Profile of Campus Ministers

Ministry Group	Number	Percent
Jewish Student Union	19	6.9
Southern Baptist Campus Ministry	54	19.6
Campus Crusade for Christ	34	12.3
The Navigators	13	4.7
Catholic Campus Ministry	36	13.0
Lutheran Campus Ministry	21	7.6
Presbyterian Campus Ministry	13	4.7
United Methodist Campus Ministry	29	10.5
Interdenominational	34	12.3
Episcopal Campus Ministry	23	8.3
Age		
Less than 33	87	31.5
33 to 42	89	32.3
43 to 52	66	23.9
Greater than 52	34	12.3
Student Body Size		
Less than 2,500 students	35	12.7
2,500 to 8,999 students	61	22.1
9,000 or more students	180	65.2
TOTAL	276	100.0

Judgment Analysis (JAN) was utilized to identify the patterns by which campus ministers make decisions about goals. The patterns were identified through the formulation of an association between the items on the instrument and an overall item. The strength of this association is reflected in the value of the multiple correlation coefficient (R). In this case the overall item represented an evaluation of all the goals which were presented to the campus ministers (Table 1). The JAN procedure gave an R^2 (multiple R coefficient squared) for each individual grouping of campus ministers and an overall R^2 for the initial stage of the procedure. The initial stage consisted of all the groupings when each one is treated as an individual system. Two judgmental groups were then selected by the procedure and combined on the basis of the homogeneity of their prediction equations. This resulted in the least loss in predictive efficiency of the procedure. The loss in predictive efficiency was measured by the drop in R^2 between the two stages. The grouping continued until all of the groupings were combined into a single cluster.

A determination of the number of different judgmental groups that are present can be made on the basis of the drop in R^2 at the different stages of the JAN procedure. Ward (1962) and Ward and Hook (1963) suggested that a drop greater than .05 between successive stages represented too great a loss in predictability.

Results

Mean responses of the 276 campus ministers are shown in Table 3. Goals which were rated as most important were number 7 (invest in

MEAN RESPONSE SCORES FOR GOAL STATEMENTS*

Category	Goal Statements																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Ministry Group:																		
1. Jewish	2.7	4.3	4.0	4.3	2.5	2.6	4.0	3.5	4.6	4.0	2.8	3.7	4.1	4.4	4.3	4.0	3.2	4.2
2. Southern Baptist	4.8	4.3	3.7	4.3	4.3	3.9	4.4	3.0	4.0	3.6	4.5	4.5	4.5	4.3	3.9	4.4	4.5	4.6
3. Campus Crusade for Christ	4.8	4.1	2.5	4.8	4.2	4.2	4.8	1.8	3.5	2.4	5.0	4.7	4.8	4.4	2.4	4.6	4.9	4.5
4. Navigators	4.7	3.8	1.6	4.2	3.8	3.2	4.8	1.8	4.1	2.7	4.9	4.5	4.8	3.6	2.7	4.4	4.8	4.2
5. Catholic	4.4	3.8	4.4	3.9	3.3	3.4	4.5	3.4	4.2	3.9	3.9	4.4	4.4	4.3	4.2	4.1	3.9	4.5
6. Lutheran	4.0	3.9	4.2	3.8	3.3	3.7	4.3	3.3	4.2	3.6	3.8	4.3	4.2	3.7	4.2	4.4	4.0	4.4
7. Presbyterian	4.0	3.7	4.2	3.8	3.1	3.7	4.4	3.9	4.3	3.5	3.6	4.3	4.2	4.2	4.5	4.5	4.1	4.3
8. United Methodist	4.0	3.6	3.6	3.8	3.3	3.7	4.2	3.2	4.2	2.9	3.7	4.1	4.2	3.9	3.7	4.3	4.0	4.4
9. Interdenominational	4.0	3.6	3.4	3.9	3.6	3.6	4.6	3.9	4.2	3.2	3.9	4.5	4.2	3.8	4.1	4.4	3.8	4.4
10. Episcopal	3.6	4.0	4.4	3.8	3.3	3.0	4.6	3.7	4.4	3.8	3.9	4.6	4.3	4.4	4.3	4.6	4.4	4.4
Minister's Age:																		
< 33 years	4.5	4.2	3.2	3.5	3.9	3.9	4.7	2.5	4.0	3.0	4.6	4.5	4.6	4.3	3.3	4.4	4.6	4.5
33-42 years	4.1	3.9	3.8	4.0	3.5	3.4	4.4	3.3	4.2	3.5	4.0	4.4	4.4	4.2	4.1	4.3	4.0	4.3
43-52 years	4.2	3.9	4.1	4.0	3.4	3.7	4.4	3.7	4.3	3.6	3.9	4.5	4.3	4.0	4.1	4.4	4.0	4.5
> 52 years	3.5	3.6	3.4	3.6	3.4	3.2	4.3	3.5	4.0	3.3	3.2	3.9	3.9	3.8	4.2	4.2	3.9	4.3
Student Body Size:																		
< 2,500 students	3.9	3.8	3.9	4.1	3.6	3.3	4.4	3.1	3.9	3.4	4.0	4.0	4.3	4.0	3.8	4.3	4.1	4.4
2,500-8,999 students	4.2	4.1	3.6	4.1	3.7	3.8	4.5	3.2	4.3	3.3	4.1	4.5	4.4	4.2	4.0	4.5	4.1	4.5
≥ 9,000 students	4.3	3.9	3.6	4.1	3.5	3.6	4.5	3.2	4.1	3.4	4.1	4.4	4.4	4.1	3.8	4.3	4.2	4.4
OVERALL	4.2	3.9	3.6	4.1	3.6	3.6	4.5	3.1	4.1	3.4	4.1	4.4	4.5	4.1	3.8	4.4	4.2	4.4

*Rounded to nearest tenth

hurting world) and number 13 (environment for growth). Each of these received an overall mean rating of 4.5. The campus ministers rated goal number 8 (organize for study and action) as having the lowest priority with an overall mean of 3.1. The Campus Crusade for Christ campus ministers gave as high as or the highest ratings of all groups for 10 of the 17 goals. The Jewish campus ministers gave as low as or the lowest ratings of all groups for 11 of the 17 goals.

Twelve of the 18 goals were given as high or the highest ratings of importance by the youngest group of campus ministers. The oldest ministers held the highest rating for only one goal, number 15 (pastoral counseling). Indeed, the oldest campus ministers had as low or the lowest ratings for 13 of the 18 goals.

Fourteen of eighteen goals were rated as high or higher by these ministers from medium sized schools than by either the ministers from schools with small or large student bodies. The small school ministers rated only one goal higher than the other two groups. That goal was number 3, i.e., to provide worship opportunities on campus.

In an effort to determine the goal orientations of the three classifications; i.e., ministry group, student body size, and age, the data were submitted to Judgment Analysis technique (JAN). Characteristics of the campus ministers who evaluated ministry goals were illuminated by JAN which incorporates the strength of association between the ratings of the 17 individual ministry goals and the overall goal rating.

Table 4 demonstrates the judgment analysis system of regrouping

Table 4

JAN ITERATIONS BY CLASSIFICATION

	Stage	Judge										R^2
MINISTRY GROUP	I	1	2	3	4	5	6	7	8	9	10	.80
	II	1	2	3	4	5	(6, 7)	8	9	10		.80
	III	1	2	3	4	(5, 8)	(6, 7)	9	10			.78
	IV	1	2	3	4	(5, 8)	(6, 7, 10)	9				.76
	V	1	(2, 5, 8)	3	4	(6, 7, 10)	9					.74
	VI	1	(2, 5, 8)	(3, 9)	4	(6, 7, 10)						.67
	VII	(1, 3, 9)	(2, 5, 8)	4	(6, 7, 10)							.59
	VIII	(1, 3, 9)	(2, 5, 6, 7, 8, 10)	4								.49
	IX	(1, 3, 4, 9)	(2, 5, 6, 7, 8, 10)									.35
	X	(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)										.19
MINISTER AGE	I	1	2	3	4							.47
	II	(1, 3)	2	4								.41
	III	(1, 3)	(2, 4)									.32
	IV	(1, 2, 3, 4)										.20
STUDENT BODY SIZE	I	1	2	3								.31
	II	(1, 2)	3									.29
	III	(1, 2, 3)										.23

classifications of ministers. This process determines the groupings who have made similar patterns in evaluating the goals. Thus in the first part of Table 4 the goals are analyzed by ministry groups. Starting with ten groupings of ministers the JAN procedure shows that group 6 (Lutheran) and group 7 (Presbyterian) were the most alike in the way that the ratings of the 17 individual goals related to the overall goal. This combination of campus ministers produced a negligibly small reduction in $\underline{R}^2$ from stage 1 to stage 2. The $\underline{R}^2$ indicated the association between the 17 goals and the overall goal for each iteration. That is, the $\underline{R}^2$ of .80 indicated that 80% of the variability in the evaluation of the overall goals was accounted for by the 17 individual goals. The iteration process continued to combine ministry groups until a .05 decline in the $\underline{R}^2$ was noted. At this time six different groupings of campus ministries out of the original ten were revealed. Groups 1, 3, 4, and 9 are singletons having distinct characteristics by themselves, whereas 2, 5, and 8 were merged and 6, 7, and 10 were merged owing to the homogeneity of their rating policies.

Using age as a means of classifying campus ministers (the second part of Table 4) four distinct ways of perceiving the subsidy of the individual goals to the overall goal of campus ministry appeared. The third part of Table 4 shows the campus ministers to have two composite policies with respect to student body size. Those campus ministers from small and intermediate size student bodies tended to have the same viewpoint concerning the contribution of individual goals to the overall while those from the largest schools were significantly different.

Discussion

The Southern Baptist, Catholic, and Methodist groups seemed to perceive all of the items as moderately associated with the overall goal of campus ministry. The goal showing the greatest contribution was number 6 (religious vocation) followed by 7 (sharing of faith).

Another composite of ministry groups combined Lutheran, Presbyterian, and Episcopal who also showed moderation on goal statements. The Presbyterians perceived goals 11 (personal ministry), 14 (visible community of faith), and 16 (relate faith) as being the most worthy dimensions of a campus ministry endeavor. While the Lutherans were very high on goals 7 (invest in a hurting world), 9 (religious identity), and 14 (visible community of faith), the Episcopals were very high on 16 (relate faith).

The other four campus ministry groups, the Jews, Navigators, Campus Crusade for Christ, and the Interdenominationals, all had very different perceptions of what a campus ministry should be. The Jews showed negative perceptions of goals 2 (fellowship), 4 (student leadership), and 6 (religious vocation) followed by negative perceptions of 10-13 (study of religious topics, personal ministry, faith development, environment for growth). All other goals seemed to make no contribution to the overall according to the perception of the Jewish ministry group. According to the Navigators goal 6 (religious vocation) has the highest priority followed by 5 (involved in local church), 4 (student leadership), and 14 (visible community of faith) for inclusion in a campus ministry program, whereas, goal 15 (pastoral counseling) was definitely not desired as a facet of a ministry program. The Campus Crusade for Christ group had high perceptions for goals 5 (involved in local church), 10 (study of religious topics), and 15 (pastoral counseling) as being foundations of a campus ministry program, whereas, the interdenominational group showed high interest in goals 15 (pastoral counseling), 16 (relate faith), and 17 (sharing of faith). The interdenominational group showed little interest in the other goals in defining their campus ministry except for number 2 (fellowship) which they perceived as not being a part of a program.

When the campus ministers were grouped by age the older personnel showed the strongest feelings about the components of a ministry program. They perceived the "lynch pins" to be composed primarily of goals 7 (invest in a hurting world), 11 (personal ministry), 13 (environment for growth), 15 (pastoral counseling), 16 (relate faith), and 17 (sharing of faith). The two middle aged groups (33 to 42 and 43 to 52) showed rather modest priority on most of the goals. The youngest of the campus ministers, however, perceived goal number 5 (involved in local church) as highest priority in a program followed by 1 (Biblically based life goals) and 14 (visible community of faith).

In the grouping according to campus population, ministers employed at small and intermediate sized campuses tended to have similar perceptions concerning the constituents of a campus ministry program. They also seemed to have the strongest perceptions overall, particularly wherein they rated goals 2 (fellowship), 3 (worship), and 7 (invest in a hurting world) as not being a part of the campus ministry goal. However, these ministers rated goals 1 (Biblically based life goals), 16 (relate faith), and 17 (sharing of faith) as being most contributory. Campus ministers from larger campuses tended to be very moderate across the board, that is, they viewed all goals as being moderately contributory to an overall campus ministry goal.

Conclusions

The study seems to have revealed a consensus of priorities concerning the components of a campus ministry mission. These components are revealed according to ministry group, age of the campus ministers, and size of the student body at the institution where the campus ministers are employed. Evidence indicates that Southern Baptists, Catholics, and Methodists dominate the campus ministry movement. They revealed a moderation concerning the components of the campus ministry mission and seemed to view the campus ministry as an extension of the affiliated institution of higher learning. Evidence further suggests that Lutherans, Presbyterians, and Episcopal campus ministers viewed the goals from the standpoint of a more orthodox form of protestantism. The literature seems to indicate that these divisions tend to have more rituals and liturgy in their activities. The Lutherans seemed to view the campus ministry as a church functioning as a community within the campus, whereas the Presbyterians tended to emphasize the importance of personal faith in campus ministry. The Episcopals on the other hand seemed to underscore the idea that the campus ministry mission should support an applied religious philosophy.

That is, religion should address questions dealing with the way one should live in contemporary times and how one should decide about situational ethics.

The Jews seemed to perceive very little social context within their campus ministry commission. They viewed the charge very differently from all other groups. Information suggests a sort of introspection about their approach. They were interested in pastoral counseling, individual religious identity and local church involvement in their campus ministry mission.

Church involvement in the student's life appeared to be a cornerstone of the Campus Crusade's ministry. The Navigators seemed to emphasize a religious leadership orientation with a social context. Results also suggest the Navigators as being organizers of leadership development. The Interdenominational group stressed individual student growth and sharing faith with other individuals.

When the sample was reclassified according to campus population, those campus ministers from small and intermediate size campuses seemed more interested in individual aspects of religious manifestations. Moreover, they were somewhat negative on fellowship and group worship. Ministers from the largest campuses seemed more attentive to social programming but were moderately involved in all 17 of the goals.

Although the lack of a clear understanding of the role of campus ministry may be a problem in the field, it can be assumed that the campus ministers participating in the present study had definite judgmental policies of campus ministry and were consistent in expressing them.

References

Flanagan, J. C. (1954). The critical incident technique. Psychological Bulletin, 51(4), 327-358.

Hammond, P. C. (1979). A response. In J. H. Westhoff (Ed.), The church's ministry in higher education. Lafayette, IN: National institute for Campus Ministries. (ERIC Document Reproduction Service No. ED 158 690)

Johnson, R. L. (1979). Ministry in secular colleges and universities. In J. H. Westhoff (Ed.), The church's ministry in higher education. Lafayette, IN: National Institute for Campus Ministries. (ERIC Document Reproduction Service No. ED 158 690)

Lanagan, M. A. (1980). The distribution of Catholic campus ministries' time among work components established by the United States Catholic Conference and the National Catholic Educational Association. Dissertation Abstracts International, 40, 5749A. (University Microfilms No. 80-10696)

Underwood, D. (1969). The church, the university, and social policy (2 vols). Middletown, CT: Wesleyan University Press.

Ward, J. H., Jr. (1962). Multiple linear regression models. In H. Borko (Ed.), Computer applications in the behavior sciences. Englewood Cliffs, NJ: Prentice-Hall.

Ward, J. H., & Hook, M. E. (1963). Application of an hierarchial grouping procedure to a problem of grouping profiles. Educational and Psychological Measurement, 23(1), 69-81.

CALL FOR PAPERS

Call for Papers:

Members and prospective members of the SIG are encouraged to submit paper proposals for SIG at the 1985 AERA meeting in Chicago. In particular, we would like to encourage "users" of regression to consider submitting proposals. We would like to have one full session on regression applications only. If you are a user of regression and would consider submitting a paper, please do so. If you feel reluctant because you feel you need additional technical expertise, we'll even try to find a "regressor" to help. Also, we should have sufficient time available to us at Chicago. We would be able to accommodate several more sessions than we had at New Orleans. Send your paper proposals to me:

John D. Williams
Professor, Measurement and Statistics
Center for Teaching and Learning
University of North Dakota
Box 8158, University Station
Grand Forks, ND 58202-8158

Also, the bottom of this sheet can be used for becoming a member of the SIG. Send your $5.00 to:

Dr. Isadore Newman
Research and Design Consultant
College of Education
The University of Akron
Akron, Ohio 44325

Name: ______________________________

Address: ______________________________

____

Telephone: ______________________________

If you are submitting a research article other than notes or comments, I would like to suggest that you use the following format if possible:

Title
Author and affiliation
Indented abstract (entire manuscript should be single 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.)

It is the policy of the M.L.R. SIG-multiple linear regression and of *Viewpoints* to consider articles for publication which deal with the theory and the application of multiple linear regression. Manuscripts should be submitted to the editor as original, double-spaced, *camera-ready copy.* Citations, tables, figures, and references should conform to the guidelines published in the most recent edition of the *APA Publication Manual* with the exception that figures and tables should be put into the body of the paper. 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, *Viewpoints* is 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.

"Membership in the Multiple Linear Regression Special Interest Group is renewed yearly at the time of the American Educational Research Association convention. Membership dues pay for a subscription to the *Viewpoints* and are either Individual at a rate of $5, or Institutional (libraries and other agencies) at a rate of $18. Membership dues and subscription requests should be sent to the executive secretary of the M.L.R. SIG."

THE UNIVERSITY OF AKRON
AKRON, OHIO 44325

AKRON
JUL 24 '84
OHIO
U.S. POSTAGE
63
P.B. METER
663282

146ROGE5
ROGERS, BRUCE G.
DEPT OG ED PSYCH
UNIV OF NO IOWA
CEDAR FALLS, IOWA 50613

BOOKS SPECIAL 4th CLASS RATE

TABLE OF CONTENTS

ISSN 0195-7171