DOCUMENT RESUME

ED 166 206 TM 007 745

AUTHOR Smith, Nick L.; Murray, Stephen L.
TITLE The Use of Path Analysis in Program Evaluation. No. 12.
INSTITUTION Northwest Regional Educational Lab., Portland, Oreg. Office of Research and Evaluation Services.
PUB DATE Feb 78
NOTE 32p.
AVAILABLE FROM Northwest Regional Educational Laboratory, 710 S.W. Second Avenue, Portland, Oregon 97204 (ISBN Number 0-89354-811-1, $3.00)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Correlation; *Critical Path Method; Data Analysis; Educational Research; *Evaluation Methods; Intermediate Grades; Mathematical Models; *Multiple Regression Analysis; Predictor Variables; Program Evaluation; *Research Design; Research Methodology

ABSTRACT

Path analysis, a technique related to multiple regression analysis is used for ascribing causal relationships among variables. Path analysis involves the construction of explicitly formulated causal models and makes the reasoning explicit in the form of path diagrams and structural equations. Regression analysis is then used to construct path coefficients or beta weights. Several assumptions are required by this model of path analysis. In the given example; teacher age, teacher sex, teacher training, and class size are assumed to influence democratic classroom control directly, and also through their influence on teacher orientation toward task accomplishment (which is assumed to influence democratic classroom control directly). There are a number of advantages of path analysis, as well as a number of technical and practical problems. Path analysis enables evaluators to conduct local decision-oriented studies and at the same time to investigate theoretical causal mechanisms. However, there are many difficult problems associated with the valid use of the technique, and care must be taken in its application. (Author/CTM)

# Research Evaluation Development

# Paper Series





Office of Research and Evaluation Services

Northwest Regional Educational Laboratory

# THE USE OF PATH ANALYSIS IN PROGRAM EVALUATION

No. 12

Nick L. Smith
Stephen L. Murray

Northwest Regional Educational Laboratory

February 1978

ISBN 0-89354-811-1

The Research, Evaluation and Development (RED) Paper Series is published by the Northwest Regional Educational Laboratory (NWREL). It presents papers dealing with both theoretical and applied aspects of research, evaluation or program development. Authors, within or outside of NWREL, wishing to submit articles for possible publication in the RED series may contact the editor, RED Paper Series, at NWREL. Persons wishing to order current or back copies of the RED series may contact the Director of Marketing. A price list for titles in the series is located at the end of this paper.

Northwest Regional Educational Laboratory
710 S. W. Second Avenue
Portland, Oregon 97204

## ABSTRACT

# The Use of Path Analysis in Program Evaluation

Nick L. Smith
Stephen L. Murray

Northwest Regional Educational Laboratory

Path Analysis techniques provide evaluative researchers with a unique capability for studying the presumed causal relationships underlying social and educational programs. With appropriate designs, it enables evaluators to conduct local decision-oriented studies while simultaneously investigating fundamental causal mechanisms. In this paper we provide a brief overview of Path Analysis and illustrate its use in an actual evaluation study. In the concluding pages, we highlight the major strengths and weaknesses of using Path Analysis in program evaluation.

TABLE OF CONTENTS

Content | Page

THE USE OF PATH ANALYSIS IN PROGRAM EVALUATION

## PATH ANALYSIS: A BASIC DESCRIPTION

As Rossi and Wright (1977) have recently noted:

> There is almost universal agreement among evaluation researchers that the randomized controlled experiment is the ideal model for evaluating the effectiveness of a public policy. If there is a Bible for evaluation, the Scriptures have been written by Campbell and Stanley (1966), along with a revised version by Cook and Campbell (1975). The "gospel" of these popular texts is that all research designs can be compared more or less unfavorably to randomized controlled experiments, departures from which are subject to varying combinations of threats to internal and external validity. (p. 13)

It is also widely recognized among evaluation researchers, however, that randomized controlled experiments can rarely be conducted except with specialized types of treatments or programs. Many conditions, such as the following, have forced evaluators away from experimental and even non-experimental approaches:

- the inability to operationalize treatment variables because of their pluralistic and political origins;
- difficulty in adequately monitoring, let alone controlling, the administration of treatment conditions due to their frequent complexity and ambiguity;
- difficulty in establishing meaningful control groups since withholding of treatment is frequently seen as illegal, unethical, or discriminatory; and
- the inability to spend the amount of time necessary to adequately conduct and analyze suitable field experiments because of immediate needs for policy information.

To date, most program evaluation studies have employed quasi-experimental cross-sectional studies without randomized control groups. Such cross-sectional studies have the advantages of providing for:

- the relatively efficient collection of data within a short time frame;
- the study of interrelationships between variables at one point in time;
- the generalization of results to populations of interest (given proper sampling);
- the elimination or control of some confounding factors; and
- the use of a vast array of analysis methods that have been developed over the years for use with cross-sectional data.

Cross-sectional studies result in static models, however, that do not provide a basis for understanding the structural or causal relationships among the components of the program being evaluated. Although causality continues to be a controversial topic in philosophy as well as in the social sciences (Broudy et al., 1973), it is a prime concern of not only the practitioners but especially the consumers of evaluation research.

Path Analysis provides a flexible means of addressing the causal problems that arise within the context of program evaluation. While path analytic techniques have been designed primarily for non-experimental data, they can also be applied to experimental, cross-sectional, and longitudinal data. Although causal inference is particularly difficult in non-experimental research, one of the virtues of Path Analysis is that in order for it to be validly applied, the researcher must make explicit both the theoretical formulations of causal relationships within the program and any associated assumptions. Path Analysis, then, provides a means of empirically testing the tenability and consistency of the assumed relationships. It forces a sharpening and testing of both the logical and empirical basis underlying programmatic causal claims.

The next secti of this paper provides a brief overview of Path Analysis. Following sections contain an application of Path Analysis in evaluating an educational program and a more detailed discussion of the advantages and problems of using Path Analysis in program evaluation studies.

## Overview of Path Analysis

Path Analysis originated in genetics with the work of Wright (1921, 1934, 1954, 1960a, 1960b) who developed the method as:

> an extension of the usual verbal interpretation of statistics, not of the statistics themselves. It is usually easy to give a plausible interpretation of any significant statistic taken by itself. The purpose of path analysis is to determine whether a proposed set of interpretations is consistent throughout. (1960b, p. 444)

Social scientists were introducted to Path Analysis primarily as a result of Blalock's (1964, 1968a, 1968b, 1971) and Duncan's (1966) use of causal modeling in sociology. Land (1969) helped to generalize the method by developing the methodology of Path Analysis from basic definitions and assumptions rather than from emphasizing the construction of basic equations within sociological theory. General discussions of Path Analysis are provided in Duncan (1966), Land (1969), Blalock (1971), Werts and Linn (1971), and Spaeth (1975). Since their introduction, Path Analysis techniques have been used in research studies on a wide variety of problems including:

job performance (Greene, 1973, Lawler, 1968, Wanous, 1974)
role perceptions (Miles, 1975)
aggression (Eron, Huesman, Lefkowitz and Wolder, 1972)
life satisfaction (Sears, 1977)
social stratification (Blau and Duncan, 1967)
educational aspirations (Sewell and Shah, 1968)
educational achievement (Anderson and Evans, 1974)
teacher characteristics (Anderson and Evans, 1974)
college environments (Werts, 1967)

Path Analysis, as most of the other forms of causal modeling, can be characterized as:

a. relating primarily to the analysis of non-experimental data and the absence of laboratory or experimental controls;

b. employing hypothetical constructs to specify presumed but latent variables which, though not always directly observed, are implied in the relationships with the observable variables; and

c. utilizing a systems orientation reflected in the use of sets of interacting relational equations (Goldberger and Duncan, 1973).

Path Analysis involves the construction of explicitly formulated alternative structural (causal) models which imply different patterns of relationships among variables. Path Analysis makes the underlying causal reasoning explicit in the form of path diagrams and structural equations. Regression analysis is then used to construct "path coefficients" (beta weights). Models inconsistent with the data are rejected, while those not rejected are viewed as plausible causal patterns to be further studied. Causation cannot be unambiguously demonstrated using these techniques, but some causal patterns can be rendered more believable than others.

Spaeth (1975) has pointed out that Path Analysis is not a single statistical procedure:

> Instead, it is a family of ways of analyzing data. The members of this family are models depicting the influence of one set of variables on another. (p. 53)

Models which may be given path analytic interpretations, and are therefore members of this family, include ordinary multiple regression, confirmatory factor analysis, canonical analysis, and two-stage least squares. It should be emphasized, therefore, that Path Analysis is not just a special use of ordinary regression analysis, though regression analysis is used to compute path coefficients. In Path Analysis, an equation represents a causal link and not just an empirical association, whereas in least squares regression, an equation simply represents the mean of the dependent variable as a function of the independent variables. Least squares regression procedures give inappropriate estimates whenever the causal equations contain unobservable variables (errors of measurement), reciprocal causation (simultaneity) or omitted variables (inadequate control). If none of these conditions exist (or at least one makes such an assumption) then least squares regression does give suitable parameter estimates. (See Goldberger and Duncan, 1973, for a detailed discussion of this point.)

The following assumptions are made in the use of Path Analysis. These assumptions have to be met to obtain valid results from an application of the technique.

1. A clearly defined, explicitly specified causal system which includes all relevant variables, is assumed.

2. The form of the model must be correctly specified with variables ordered correctly from the point of view of both theory and measurement procedures.

3. Within the structural model a change in one variable is always a linear function of changes in other variables.

4. Dependent variables are assumed to be uncorrelated with each other.

5. Hypothetical, unmeasured variables are assumed to be continuous.

6. Measurements are assumed to have a high degree of reliability and validity (low measurement error).

7. The data must meet all multivariate regression assumptions including:

- interval scale measurement,
- homoscedasticity,
- relatively low intercorrelations among causal variables used as predictors,
- linear and additive effects among variables,
- residuals are uncorrelated with a mean of zero and unit variance,
- fixed effects independent variables.

Path models may be recursive (involving only one-way causation) or non-recursive (involving reciprocal causation). For the sake of clarity and simplicity, this paper deals only with recursive models.

A path diagram is a graphic display of the order in which variables are assumed to affect one another. Variables in a path diagram are either endogenous or exogenous. An endogenous variable is one which is dependent upon other variables in the diagram while exogenous variables are those which affect endogenous variables. Since there can be more than one stage of causation displayed in a path diagram, some endogenous variables may act as exogenous variables for subsequent endogenous variables. The initial variables in a diagram, however, are always exogenous to that system.

Conventions for drawing path diagrams are described by Land (1969). Arrows are drawn from variables acting as causes to variables acting as effects. Initial exogenous variables are linked to one another by curved lines with double arrows. For example, suppose one assumes that X is caused by both A and B and that one is not interested in the causes of A or B. This would be represented as below in Figure 1.

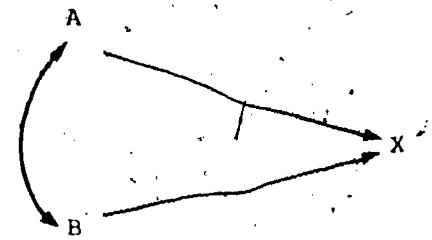


Figure 1. A simple causal model

Analysis is designed to produce measures of relationship between a given endogenous variable and each of the exogenous variables on which it is dependent. These indices are called "path coefficients." The reduction of the effects of any exogenous variable into its <u>direct effects</u> (direct paths of influence), and its <u>indirect effects</u> (effects analogous to the effects of an antecedent variable as transmitted through an intervening variable), is a major feature of the Path Analysis technique.

The reader is referred to Land (196[illegible]) for an extensive discussion of the principles of Path Analysis including derivations and interpretations of path coefficients in various types of path models. Spaeth (1975) provides a detailed comparison of the logics behind predictive least squares models and totally recursive path models.

## PATH ANALYSIS: AN APPLICATION

Before proceeding to a more detailed discussion of the general advantages and problems of using Path Analysis in program evaluations, a specific example is offered here to help familiarize the reader with the basic approach. This example concerns the application of Path Analysis to assess the classroom impact of a teacher training workshop. (This example was one facet of a more comprehensive evaluation study that was reported in Murray <u>et al.</u>, Note 1.)

## The Evaluation Problem

An evaluation study was conducted of an instructional system designed to increase the skills of school teachers in systematically carrying out a five-step method of problem solving: (a) identify the problem, (b) diagnose the problem situation, (c) consider alternative actions, (d) try out a plan of action, and (e) adapt the plan.

Participants are guided through 16 sequential and cumulative training units, each consisting of a series of concept papers, group discussions, and exercises. The materials were designed to develop participant knowledge and participant ability to use the problem solving process in identifying and diagnosing classroom, school, and peer-related problems.

The instructional strategy of the system is based on a pattern of repeated diagnoses carried out in small training groups of three or six persons. The knowledge gained through this process provides the basis for selecting and designing action plans to solve the identified problems. A simulation exercise provides opportunities for participants to practice skills in training groups and to learn to observe and improve their teamwork behaviors in a workshop setting.

Thirty-eight fourth, fifth, and sixth grade teachers comprised the treatment group that attended instructional workshops in one of two locations. Data were also gathered from another 24 fourth, fifth, and sixth grade teachers who constituted a non-randomized control group. Analysis of a background questionnaire revealed that the treatment group and the control group were similar in terms of proportion of males to females, age, years of teaching experience, and reasons for attending the workshop. However, a somewhat higher proportion of control group members had obtained a master's degree.

One purpose of this evaluation study was to investigate the impact of workshop training on classroom activities. The pattern of relationships among outcome measures was examined through the use of Path Analysis.

Evaluating the impact of teacher training on students requires the definition of a sequence of outcome measures which allow the training activities to be linked with student outcomes. When there are well-defined student objectives, evaluating the impact of teacher training is straightforward. If student objectives are missing, as they were for this system, the evaluator must use what information is available to infer a sequence of training effects.

Such an inferred sequence of effects can be portrayed in a path diagram and Path Analysis can be used to test the inferred causal relationships. Representing treatment effects in a path diagram also helps identify "weak links" in the chain of effects expected to result from training.

One of the path diagrams used to assess classroom impact in this evaluation study is presented on the following page. Teacher Age, Teacher Sex, Class Size and Training are assumed to influence Democratic Classroom Control both directly and through their influence on Teacher Orientation Toward Task Accomplishment, which is assumed to influence Democratic Classroom Control directly.

These variables were selected for this path analysis on the basis of their compatibility with the overall rationale of the instructional system being evaluated. Teacher Age, Teacher Sex, Class Size and Training (a dummy variable where treatment group = 1 and control group = 0) were all assumed to be reliably measured.

Teacher Orientation Toward Task Accomplishment, which was administered prior to and immediately following the workshop, is one scale from a self-report measure of teacher orientation toward problem solving. In the questionnaire, respondents are presented with a variety of realistic problem situations and asked to estimate the probability of their behaving or reacting in a manner consistent with the view of problem solving advocated in the instructional materials. Responses to items are made in a five-point multiple-choice scale identical to that in the following sample item:

> Suppose something has gone wrong in your school--something that affects everyone and has everyone upset. What are the chances that you would remain quiet and wait for others to analyze the problem?
>
> 1. Almost none (less than a 10 percent chance)
> 2. Maybe (about a 25 percent chance)
> 3. Possibly (about a 50 percent chance)
> 4. Almost always (about a 75 percent chance)
> 5. Definitely (greater than a 90 percent chance)

A split-half reliabilty of .71, corrected with the Spearman Brown Prophecy formula, was obtained for Teacher Orientation Toward Task Accomplishment based on a sample of 87 teachers. A test-retest reliability of .61 was obtained on this scale for the same group of 87 teachers. Since the

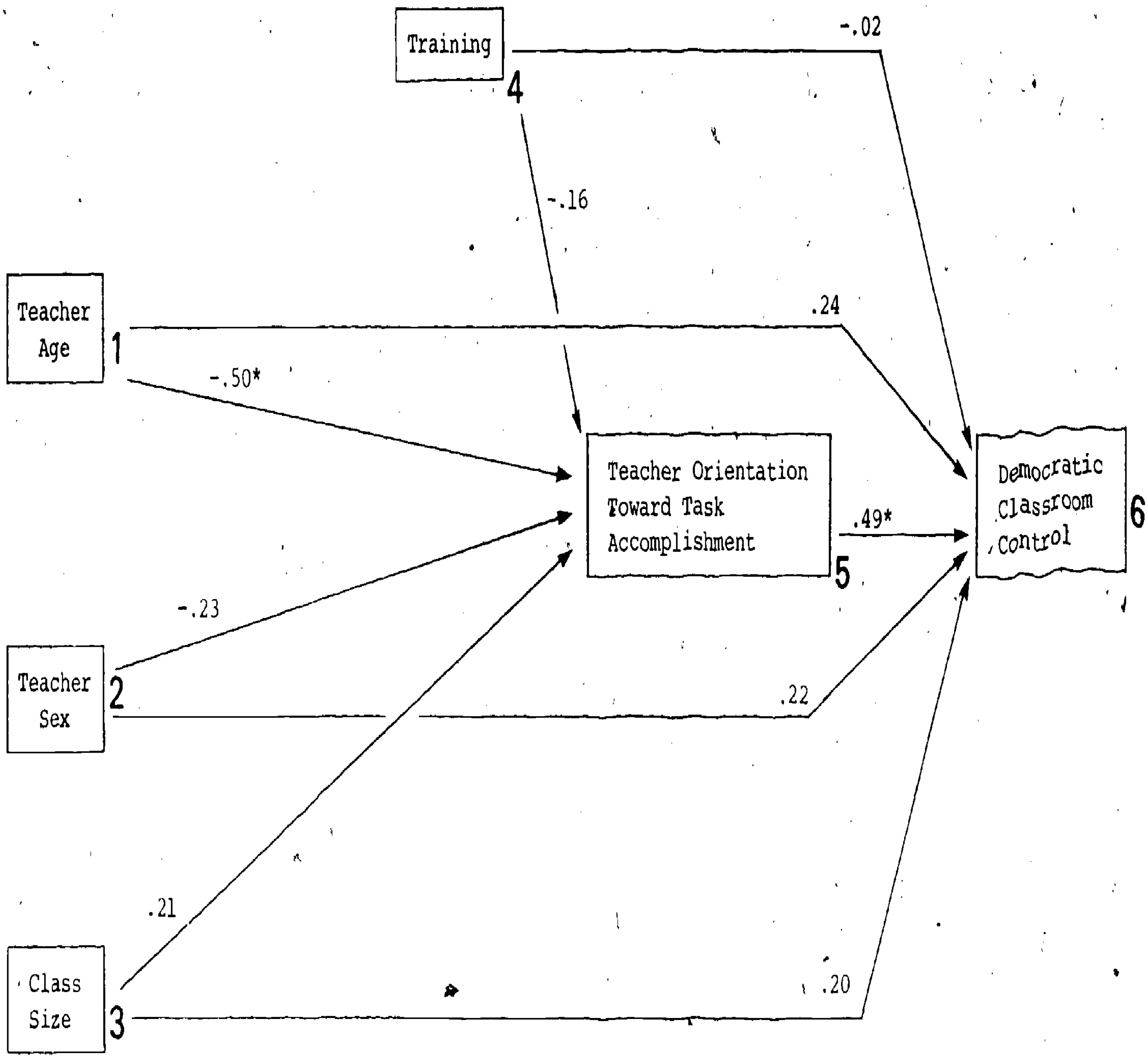


Figure 2. Path diagram and standardized partial regression coefficients for the evaluation of teacher training in terms of its impact on the level of democratic classroom control

*$p \leq .01$

test-retest reliability was based upon pretest-posttest administrations designed to assess training effects, the test-retest reliabilities should be interpreted as conservative estimates of stability.

Democratic Classroom Control is one scale on the Student Activities Questionnaire (Ellison, Callner, Fox, and Taylor, Note 2), a multiple-choice questionnaire designed to measure various aspects of classroom climate. It was administered to the treatment group after the first week of workshop instruction and again after the completion of training three months later. Similar data were collected from the control group, although due to scheduling difficulties, only two months elapsed between the pre- and posttest administrations. This scale is a measure of student input into classroom decision making, planning of individual activities, and enforcement of rules. Students with high scores on this scale report that they are allowed frequent input through discussions and planning activities and that decisions are made through the joint effort of teacher and students. This scale was selected because of its compatibility with the "group problem-solving" orientation of the instructional system and its compatibility with information on how past trainees had used the training in their classrooms.

Intraclass correlations of .79 on the pretest (based on 84 classrooms, 1,499 students), and .72 on the posttest (based on 73 classrooms, 1,213 students) were obtained for Democratic Classroom Control. (The intraclass correlation is a function of within-group and between-group variation of student responses and constitutes a measure of consistency or reliability using the classroom as the unit of analysis.) A test-retest reliability of .45 was also obtained (based on 73 classrooms, 1,213 students) but must be considered as a conservative estimate because of the intervening treatment.

The Path Model

Two recursive structural equations correspond to the path diagram represented in Figure 2.

$$Z_5 = P_{5,1}Z_1 + P_{5,2}Z_2 + P_{5,3}Z_3 + P_{5,4}Z_4 + P_{5,a}Z_a$$

$$Z_6 + P_{6,1}Z_1 + P_{6,2}Z_2 + P_{6,3}Z_3 + P_{6,4}Z_4 + P_{6,5}Z_5 + P_{6,b}Z_b$$

The first equation represents the influence of the exogenous variables on Teacher Orientation Toward Task Accomplishment. By solving this linear regression equation it is possible to make estimates of the direct effects (path coefficients) of the prior variables on this endogenous variable. While these path coefficients do not in themselves indicate causality they can be used to assess whether or not the data are consistent with the prior causal model. Table 1 presents the standardized partial regression coefficients or path coefficients for the variables in Figure 2 and the multiple correlation obtained for explaining post-training Teacher Orientation Toward Task Accomplishment.

Table 1

Multiple Correlation and Standardized Partial Regression Coefficients for Explaining Posttraining Teacher Orientation Toward Task Accomplishment[a]

| Exogenous Variable | Multiple Correlation | Standardized Partial Regression Coefficient | p-Value |
|---|---|---|---|
| Teacher Age | .46 | -.50 | .01 |
| Teacher Sex | | -.23 | .19 |
| Class Size | | .21 | .23 |
| Training | | -.16 | .37 |

[a] Only subjects for whom there were complete data were used in this analysis. (Treatment Teachers N=26, Control Teachers N=10)

It is apparent from examining Table 1 that the training did not have any direct effect on Teacher Orientation Toward the Task Accomplishment as the corresponding coefficient was -.16. The multiple correlation resulting from regressing posttraining Teacher Orientation Toward Task Accomplishment on the four variables in the path model was .46. The only significant path coefficient was for Teacher Age ($p \leq .01$), indicating that younger teachers saw themselves as having stronger orientations toward problem solving.

The second structural equation:

$$Z_6 = P_{6,1}Z_1 + P_{6,2}Z_2 + P_{6,3}Z_3 + P_{6,4}Z_4 + P_{6,5}Z_5 + P_{6,b}Z_b$$

represents the influences of all five variables on the classroom impact variable, Democratic Classroom Control. Table 2 includes the multiple correlation and the path coefficients indicating the relationships between each of the exogenous variables and the level of Democratic Classroom Control.

Table 2

Multiple Correlation and Standardized Partial Regression Coefficients for Explaining Posttraining Level of Democratic Classroom Control[a]

| Exogenous Variable | Multiple Correlation | Standardized Partial Regression Coefficient | p-Value |
|---|---|---|---|
| Teacher Age | .57 | .24 | .21 |
| Teacher Sex | | .22 | .20 |
| Class Size | | .20 | .24 |
| Training | | -.02 | .93 |
| Posttraining Teacher Orientation Toward Task Accomplishment | | .49 | .01 |

[a]Only subjects for whom there were complete data were used in this analysis. (Treatment Classrooms N=26, Students N=650; Control Classrooms N=10, Students N=281)

The multiple correlation of .57 is significant ($p \leq .03$) and indicates that the model explains 32.5 percent of the variation in the classroom impact variable. The path coefficients (standardized partial regression coefficients) indicate that the data supported the hypothesis that teacher problem orientation (Teacher Orientation Toward Task Accomplishment) influenced classroom climate as perceived by students (Democratic Classroom Control). However, there was no evidence that training had any impact on classroom climate.

In summary, the analysis of the model represented in Figure 2 suggests that:

- younger teachers saw themselves as having stronger orientations towards task accomplishment;
- post-training teacher orientation towards task accomplishment influenced subsequent classroom climate as perceived by students; but
- training had no influence on either task orientation or classroom climate.

An analysis of covariance, which controlled for group pretest differences, had been conducted prior to the causal analysis and no significant training effect had been found. By using Path Analysis with the non-randomized control group, it was possible to not only test for direct and indirect training effects, but to investigate the causal relationships which were assumed to lead to the desired classroom impact. The causal model supported by the data would look like Figure 2 without the two training arrows.

Having briefly illustrated one means of using Path Analysis in an evaluation study, a more complete discussion of the general advantages and problems associated with the use of the technique in program evaluation is now appropriate.

## PATH ANALYSIS: AN EVALUATIVE TECHNIQUE

Path Analysis is a powerful technique that offers exceptional capabilities to the evaluative researcher. There are major difficulties, however, in its application that limit its use in program evaluation work. The strengths and problems of Path Analysis are discussed more fully here in this final section of the paper.

### Strengths of Path Analysis

Path Analysis provides the program evaluator with unique advantages and opportunities not available with most other techniques. Several of these special advantages are highlighted in the following:

Provides a means for modeling of complex programs. Path Analysis provides a means of incorporating recursive and reciprocal causation and unobservable variables into evaluative studies of programs. The resultant analytic models are a more faithful representation of the dynamic reality of the programs themselves. With this technique, programs will not need to be simplified "beyond recognition" just in order to be evaluated.

Forces consideration of alternative models. Path Analysis requires the explicit specification of presumed causal relationships and forces the researcher to consider several alternative causal models. (In the example presented above, twelve alternative causal models were identified and considered in the course of the study.) These competing models can then be tested simultaneously using not only correlational information, but other observational and experimental data. The technique, therefore, broadens the realm of inquiry and helps forestall the "pet theory" and "objectives driven" biases evident in much evaluative research. Path Analysis provides one means of preserving an evaluator's objective detachment by enabling him or her to simultaneously investigate multiple working hypotheses (Chamberlin, 1965).

Enables the study of both direct and indirect effects. Path Analysis enables the researcher to study both direct and indirect effects on dependent variables as well as to analyze correlations between variables into various components, depending on the causal models postulated. Path analytic techniques provide a means of distinguishing between (1) components due to the mediated effects of one variable through another variable, (2) components which are due to correlations between various causes in the system, and (3) spurious components which are due to one variable having a common cause with another variable in the model. The identification of indirect effects is crucial both when evaluating program impact and when making program revisions.

Permits incremental revision of the structural model. If a path model is overidentified (that is, the information from the observable variables provides more than enough information to determine the structural parameters, as when certain path coefficients are hypothesized to equal zero), then it is possible to employ this information to revise the model. As a part of the evaluative study, special testing procedures can then be employed which indicate what parts of the model need revision. These procedures allow the evaluator not only to test the tenability of the overall model, but to revise sections of the model, all within the context of one evaluative study. (See, for example, Pedhazur, 1975, and Goldberg and Duncan, 1973.)

Enables the estimate of the severity of the various biases. Sources of bias and the means of estimating their magnitude have been extensively studied for traditional least squares methods. Since Path Analysis is based on these methods, it is possible to identify and measure the sources of bias in path models to a much greater extent than with most other techniques in evaluative research. These sources of bias are identified in the sections below concerned with problems in Path Analysis.

Can be used on a heuristic, logical, or statistical level. As illustrated above, Path Analysis was designed primarily as a logical and statistical approach for testing hypothesized causal models. Path diagrams can be used, however, as a purely heuristic device for charting presumed causal relations. The explication of causal relations can be useful in planning evaluation studies (whether they are to be experimental or non-experimental in nature) and in studying the logical consistency of the implicit "theories in use" which guide program development and relate program activities to program goals.

Permits developmental testing of innovative programs. The development of innovative programs is sometimes based on explicit rationales which are contrary to standard practice and which have little substantiation in past research. Path Analysis provides a means of explicitly delineating the rationales of these innovative models and of testing their tenability in their current developmental forms.

Provides greater continuity and generalizability to evaluation findings. Path analytic techniques provide a means of studying the basic causal relationships within social and educational programs. By including treatment and comparison groups in path models, it is possible to conduct comparative evaluation studies helpful to decision makers and at the same time to conduct research studies of basic programmatic mechanisms. Such an approach provides both local decision-making information, and the more basic programmatic information that can be accumulated and generalized across succeeding evaluation studies.

Problems with Path Analysis

As Feldman (1975) points out

> Like most advances, however, causal-correlational analysis is not an unmixed blessing. To be used properly, many issues must be dealt with in each study, on both statistical and theoretical levels. Failure to consider these issues leads to serious restrictions on the inferences that may be drawn and on the theoretical relevance of a given set of results. (p. 663)

Some of the technical problems associated with the use of path analysis procedures are presented below, followed by a discussion of the more pragmatic problems. The reader is referred to the cited publications for fuller discussions of the technical problems which follow:

The Problem of Congruent versus Incongruent Influence

In a path diagram, if variable A appears to be causing changes in variable B in a congruent direction, that is, A is increasing the agreement between A and B scores, it could be instead that B is causing changes in A in an incongruent direction, that is, B is decreasing the agreement between A and B scores. Since the Path Analysis method alone does not specify which of these interpretations is accurate, there is an inherent ambiguity in the interpretation of the findings (see Rozelle and Campbell, 1969, Feldman, 1975).

The Influence of Causal Interval on the Structural Relations

In some cases the structural relations between variables in a path model differ depending on the causal interval used. For example, Sears (1977) found that in studying life satisfactions over a 50-year life cycle, the highest predictive value of any variable is for variables measured at the succeeding decade. Thereafter, a variable's direct predictive value diminishes with time. Similarly, Porter et al. (1974) found systematic changes over time in the correlation between job attitudes and turnover. Since structural relations can, therefore, differ depending on the causal interval used, a theoretical justification is required for the use of any given causal interval in a study. (See Feldman, 1975.)

The Problem of Necessity and Sufficiency

In determining whether or not A causes B, it is necessary to determine whether A is a necessary or sufficient cause of B, or both, and whether it is the presence or absence of A that exerts the causal influence. Although the design of the data collection may make such a determination possible, the mathematics of Path Analysis per se do not. (See Howard and Krause, 1970, Feldman, 1975.)

## The Problem of States versus Changes in States

Both causes and effects may be states (static) or changes in states (dynamic). With continuous variables there are four different causal patterns possible:

- static-static: A certain state of A causes a certain state of B;
- static-dynamic: A certain state of A causes a certain change in the state of B;
- dynamic-static: A certain change in the state of A causes a certain state in B;
- dynamic-dynamic: A certain change in the state of A causes a certain change in the state of B.

Finding that one pattern exists does not necessarily eliminate the other three. (See Howard and Krause, 1970, Feldman, 1975). In program evaluation, this problem is related to the use of posttest scores (static) or change scores (dynamic) in assessing program impact on dependent variables.

## The Problem of Multicollinearity

When the exogenous variables in a causal model are interrelated (multicollinearity), it is not possible to unambiguously determine what proportion of the variance in the endogenous variable is accounted for by each of the exogenous variables. If the exogenous variables are highly intercorrelated, then slight changes in the patterns of intercorrelations (due perhaps to measurement, specification or sampling errors) may result in substantial changes in the magnitudes of the regression coefficients. Multicollinearity often arises when multiple indicators are used in the regression estimates for the variables of prime interest. (See Pedhazur, 1975.)

## The Problem of Measurement Errors

One of the assumptions of Path Analysis is that the measurements have perfect reliability and validity, i.e., are free of error. The greater the errors of measurement, however, the more questionable the interpretation of the relative magnitudes of the path coefficients in the structural equations. (See Pedhazur, 1975, for a long list of references.)

## The Problem of Specification Errors

Specification errors result from the use of a wrong or inappropriate model, such as when relevant exogenous variables are omitted from the regression equation, when irrelevant variables are included, or when linear or additive models are used but are inappropriate.

Although specification errors can be controlled in experimental research through randomization, they cannot be controlled in non-experimental research and the subsequent addition or deletion of variables to the path model may substantially alter the magnitudes or signs of the regression coefficients. Only clear, strong theoretical formulations of the path models can forestall serious specification errors. (See Spaeth (1975) for an extensive discussion; also Pedhazur, 1975.)

The Problems of Nonlinearity and Nonadditivity

Path Analysis assumes both linearity and additivity. Though in some cases non-linearity and non-additivity (interaction effects) can be handled through the use of dummy variables and polynomial regression, the interpretation of the regression coefficients in these cases remains ambiguous and problematic. Though Path Analysis assumes linear additive models, such models may be too simplistic for some program evaluation applications. (See Pedhazur, 1975.)

The Problem of Parameter Identification

Because standard regression equations contain only observable variables, the values of equation parameters can be directly deduced from the observable variables. By contrast, the structural equations in Path Analysis may contain unobservable variables and the problem arises to what techniques to use to uniquely deduce the values of all the structural parameters (some of which may be unobservable) from just the observable variables. The identification of parameter values under reciprocal causation presents a problem, while models that assume one-way causation and non-correlation of residual variables are completely identifiable and so do not have this problem. (See Goldberger and Duncan, 1973, and Land, 1969).

In addition to the above technical problems, there are some pragmatic problems which may preclude the use of Path Analysis in many program evaluation studies.

Time and Resource Requirements

When Path Analysis is used in evaluations to assess treatment effects over specific causal intervals, then data must be collected on multiple variables at several points in time. Because of the complex modeling and additional data collection, Path Analysis is likely to be more expensive and time consuming than simple pre-post procedures.

A Priori Knowledge Demanded

The proper use of Path Analysis requires that the evaluator be able to a priori specify all relevant variables and their order and relationship in one or more causal models since a five-variable system can be portrayed in 1,048,576 possible configurations

(Young, 1977). Path Analysis may not be appropriate, therefore, for developmental or innovative educational programs where the presumed causal relationships are largely unknown or are unidentifiable due to continual shifts in program rationale.

### Complexity of Interpretation and Use

Path models with four or more variables become increasingly complex and problematic to use. It may be that models which appropriately represent most programs will be simply too complex and the calculation of coefficients and their interpretation will be too tedious to make the technique of practical value in many applied evaluation studies.

## Conclusion

Path Analysis provides evaluative researchers with a unique capability for studying the presumed causal relationships underlying social and educational programs. With appropriate designs, it enables evaluators to conduct local decision-oriented studies and at the same time to investigate more fundamental causal mechanisms--an excellent means of combining field research with field service.

There are many difficult problems associated with the valid use of the technique, however, and appropriate care must be taken in its application. It remains to be seen just how useful Path Analysis will be in a variety of applied evaluation settings. However, it appears to be a promising approach for program evaluation which should now be actively tested through thoughtful use.

## REFERENCE NOTES

1. Murray, S. L., Rassen, R. L., and Speedie, S. M. Research Utilizing Problem Solving: Outcome Evaluation Report. Portland, Oregon: Northwest Regional Educational Laboratory, 1976.

2. Ellison, R. L., Callner, A., Fox, D. G., and Taylor, C. W. The Measurement of Academic Climate in Elementary Schools. Paper presented at the 1973 American Psychological Association Convention, Montreal. Salt Lake City: University of Utah, 1973. (mimeograph)

REFERENCES


Anderson, J. G., and Evans, F. B. Causal models in educational research: Recursive models. American Educational Research Journal, 1974, 11 (1), 29-39.

Blalock, H. M. Causal Inferences in Non-Experimental Research. Chapel Hill: University of North Carolina Press, 1964.

Blalock, H. M. Causal Models in the Social Sciences. Chicago: Aldine, 1971.

Blalock, H. M. The measurement problem: a gap between the languages of theory and research. In H. M. Blalock and A. B. Blalock (Eds.) Methodology in Social Research. New York: McGraw-Hill, 1968(a).

Blalock, H. M. Theory building and causal inferences. In H. M. Blalock and A. B. Blalock (Eds.) Methodology in Social Research. New York: McGraw-Hill, 1968(b).

Blau, P. M. and Duncan, O. D. The American Occupational Structure. New York: Wiley, 1967.

Broudy, H. S., Ennis, R. H., and Krimerman, L. I. (Eds.). Philosophy of Educational Research. New York: John Wiley and Sons, 1973.

Campbell, D. T., and Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally, 1966.

Chamberlin, T. C. The method of multiple working hypotheses. Science, 148, 1965, 754-759.

Cook, T. C., and Campbell, D. T. The design and conduct of quasi-experimental and true experiments in field settings. In M. D. Dunnette (Ed.) Handbook of Industrial and Organizational Research. Chicago: Rand-McNally, 1975, 223-326.

Duncan, O. D. Path analysis: Sociological examples. American Journal of Sociology, 1966, 72, 1-16.

Eron, L. D., Juesmann, L. R., Lefkowitz, M. M., and Wolder, L. O. Does television violence cause aggression? American Psychologist, 1972, 27, 253-263.

Feldman, J. Considerations in the use of causal-correlational techniques in applied psychology. Journal of Applied Psychology, 1975, 60 (6), 663-670.

Goldberger, A. S. and Duncan, O. D. (Eds). Structural Equation Models in the Social Sciences, Seminar Press, New York, 1973.

Green, C. N. Causal connections among manager's merit pay, job satisfaction, and performance. Journal of Applied Psychology, 1973, 58, 95-100.

Howard, K. I. and Krause, M. S. Some comments on "Techniques for estimating the source and direction of influence in panel data". Psychological Bulletin, 1970, 74, 219-224.

Land, K. C. Principles of path analysis. In E. F. Borgatta and G. W. Bohrnstedt (Eds.) Sociological Methodology: 1969. San Francisco: Jossey-Bass, 1969, 3-37.

Lawler, E. E. A correlational-causal analysis of the relationships between expectancy attitudes and job performance. Journal of Applied Psychology, 1968, 52, 462-468.

Miles, R. H. An empirical test of causal inference between role perceptions of conflict and ambiguity and various personal outcomes. Journal of Applied Psychology, 1975, 60, 334-339.

Pedhazur, E. J. Analytic methods in studies of educational effects. In F. N. Kerlinger (Ed.) Review of Research in Education. Itasca: Illinois, Peacock, Vol. 3, 1975, 243-286.

Porter, L. W., Steers, R. M., Mowday, R. T., and Boulian, P. V. Organizational commitment, job satisfaction, and turnover among psychiatric technicians. Journal of Applied Psychology, 1974, 59, 603-609.

Rossi, R. H., and Wright, S. R. Evaluation Research: An evaluation of theory, practice, and politics. Evaluation Quarterly, 1 (1), 1977, 5-52.

Rozelle, R. M. and Campbell, D. T. More plausible rival hypotheses in the cross-lagged panel correlation technique. Psychological Bulletin, 1969, 71, 74-80.

Sears, R. R. Sources of life satisfactions of the Terman gifted man. American Psychologist, 1977, 32 (2), 119-128.

Sewell, W. H., and Shah, V. P. Social class, parental encouragement, and educational aspirations. American Journal of Sociology, 1968, 73, 559-572.

Spaeth, J. L. Path analysis. In D. J. Amick and H. J. Walberg (Eds.) Introductory Multivariate Analysis. Berkeley: McCutchan, 1975. 53-89.

Wanaus, J. P. A causal-correlational analysis of the job satisfaction and performance relationship. Journal of Applied Psychology, 1974 59, 139-144.

Werts, C. E. The study of college environments using path analysis. National Merit Scholarship Research Reports, 1967, 3, 1-40.

Werts, C. E., and Linn, R. L. Pathanalysis: Psychological examples. Psychological Bulletin, 1971, 74, 193-212.

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

Wright, S. The interpretation of multivariate systems. In O. Kempthorne et al. (Eds.) Statistics and Mathematics in Biology. Ames, Iowa. State College Press, 1954, 11-33.

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

Wright, S. Data coefficients and path regressions: alternate or complementary concepts? Biometrics, 1960(a), 16, 189-202.

Wright, S. The treatment of reciprocal interaction, with or without log, in path analysis. Biometrics, 1960(b), 16, 423-445.

Young, J. W. The function of theory in a dilemma of path analysis. Journal of Applied Psychology, 1977, 62 (1), 108-110.

The following papers in the Research, Evaluation and Development Paper Series are available at the prices stated below:

1. The Status of Research on Models of Product Development and Evaluation, by Nick L. Smith and Stephen L. Murray, Northwest Regional Educational Laboratory. 16 pp., ISBN 0-89354-800-6, $2.00

2. Techniques and Procedures for Formative Evaluation, by James Sanders, Northwest Regional Educational Laboratory and Donald J. Cunningham, Indiana University. 46 pp., ISBN 0-89354-801-4, $3.50

3. A Look at the Mosaic of Educational Evaluation and Accountability, by Blaine Worthen, Northwest Regional Educational Laboratory. 36 pp., ISBN 0-89354-802-2, $3.00

4. The Administrator-Researcher Interaction: The Conduct of Cooperative Research, by Stephen L. Murray and Nick L. Smith, Northwest Regional Educational Laboratory. 14 pp., ISBN 0-89354-803-0, $2.00

5. Some Tasks and Competencies Frequently Required in Educational Research and Evaluation, by Blaine Worthen, Northwest Regional Educational Laboratory. 26 pp., ISBN 0-89354-806-5, $3.00

6. Pre-Product Evaluation: What to do While Waiting for the Product, by Nick L. Smith, Northwest Regional Educational Laboratory. 13 pp., ISBN 0-89354-804-9, $2.00

7. Sources of Values Influencing Educational Evaluation, by Nick L. Smith, Northwest Regional Educational Laboratory. 32 pp., ISBN 0-89354-809-X, $3.00

8. A Prescriptive Model of Development or Evaluation: Some Needed Maturity, by Cecil Clark, Brigham Young University. 38 pp., ISBN 0-89354-805-7, $3.50

9. The Use of Multiple Strategies in the Evaluation of Experience-Based Career Education, by Thomas R. Owens, Joseph F. Haenn and Harry L. Fehrenbacher, Northwest Regional Educational Laboratory. 36 pp., ISBN 0-89354-807-3, $3.00

10. The Use of Student Case Study Methodology in Program Evaluation, by Harry L. Fehrenbacher, Thomas R. Owens and Joseph F. Haenn, Northwest Regional Educational Laboratory. 34 pp., ISBN 0-89354-808-1, $3.00

Continued

11. Concept Marketing of Educational Products and Services--The RDAS Approach, by Robert G. Green, Northwest Regional Educational Laboratory. 15 pp., ISBN 0-89354-810-3, $2.00

12. The Use of Path Analysis in Program Evaluation, by Nick L. Smith, and Stephen L. Murray, Northwest Regional Educational Laboratory. 25 pp., ISBN 0-89354-811-1, $3.00

Complete set of twelve titles: $28.00 (reflects 15% discount from $33.00 full price). Prices subject to change without notice.

Send purchase order or prepayment for any of the titles listed above to:

Office of Marketing
Northwest Regional Educational Laboratory
710 S.W. Second Avenue
Portland, Oregon 97204