

This is to certify that the

thesis entitled

DATA ANALYSIS STRATEGIES FOR QUASI-EXPERIMENTAL STUDIES WHERE DIFFERENTIAL GROUP AND INDIVIDUAL GROWTH RATES ARE ASSUMED presented by

STEPHEN F. OLEJNIK

has been accepted towards fulfillment of the requirements for

<u>Ph.D.</u> degree in <u>Counseling</u>, Personnel Services & Educational Psychology

Major professor

Date ______0ctober 28, 1977

O-7639

DATA ANALYSIS STRATEGIES FOR QUASI-EXPERIMENTAL STUDIES WHERE DIFFERENTIAL GROUP AND INDIVIDUAL GROWTH RATES ARE ASSUMED

Вy

Stephen F. Olejnik

.

A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Personnel Services and Educational Psychology

DATA ANALYSIS STRATEGIES FOR QUASI-EXPERIMENTAL STUDIES WHERE DIFFERENTIAL GROUP AND INDIVIDUAL GROWTH RATES ARE ASSUMED

By

Stephen F. Olejnik

Selecting an appropriate analysis strategy for a study based on a quasi-experimental research design has been a topic of considerable controversy. Recently, the discussion has focused on the issue of academic growth rates. Some authorities have suggested that the initial difference between comparison groups on a pretest measure implies that comparison groups are growing at different rates academically. The differential growth rate problem has been labeled the fan spread hypothesis. This theory suggests that along with an increasing mean difference between groups there is a proportional increase in the within group variability. Under this model traditional analyses techniques have been challenged as inappropriate on the basis that they underadjust for group differences.

The purpose of the study was to compare four analytic strategies for quasi-experimental studies where differential group and individual growth rates are assumed. Two models of growth were considered. One assumed that both individual and group growth were linear. The second model assumed linear group growth but an individual's growth rate within the group was assumed to vary over time. Under the second model of within group growth, an individual's relative academic position within

Stephen F. Olejnik

the group was likely to change. For each of these models four analytic strategies were considered: (1) gains in standard scores; (2) single covariable analysis of covariance with estimated true scores; (3) gain scores adjusted for differential growth rates; and (4) multiple fallible covariable analysis of covariance. The appropriateness of each procedure under the two models was based on two criteria: (1) the effect estimated, and (2) the precision with which each effect was estimated.

The results of the study showed that when individuals and groups grow at different rates but in a linear fashion, gains in standard scores, estimated true score analysis of covariance and gain scores adjusted for differential growth rates, all provide the correct estimate of group differences. For the second model of growth considered, gains in standard scores and gain scores adjusted for differential growth rates were both shown to estimate the desired effect. Only the multiple fallible covariate analysis of covariance procedure was shown to be an inappropriate technique for both models of differential growth.

The standard error associated with each competing analytic strategy was then compared for sample sizes ranging from 20 to 120 and the correlation between the pretest and posttest measures ranging from .1 to .9. The results indicated that when the sample is large and the correlation between the measures high, all three procedures provide approximately equal precision. As the sample size and the relationship decrease, greater precision was provided by the gain score technique adjusted for differential growth rates. The analysis also indicated that gains in standard scores provided a spuriously low standard error as a result of the two stage process of first calculating the adjusted variable and then estimating the treatment effect. As the sample size and relationship decreased, the greater the underestimation in the standard error.

It was concluded that when data are available from two points in time prior to the period investigated, adjusted gain scores provide the best analytic strategy of those considered where group growth is linear. If data are available from a single point in time and individuals and groups are growing differentially but linearly, then true score analysis of covariance provides the preferred analytic approach. In situations where only the results of a single pretest are available and individual growth rates vary across time, gains in standard scores may be appropriate but the sample size must be large. To demonstrate these results a data set was obtained and analyzed using the three competing strategies. The findings of this analysis were consistent with those predicted above.

ACKNOWLEDGMENTS

I am especially grateful to Dr. Andrew C. Porter, my friend and the chairman of my dissertation committee, for his support and guidance throughout this research project and my graduate program. By giving generously of his time and knowledge, Andy has enriched my educational experiences and made them happy ones. I would also like to thank Dr. Kenneth Arnold for his assistance, especially in the mathematical aspects of this study. Sincere appreciation is extended to Drs. John Schweitzer, Robert Flooden, and William Schmidt for their constructive criticisms and suggestions in earlier drafts of this dissertation.

I am also grateful to Ms. Janice Cave for her encouragement throughout my doctoral program and her assistance in the typing of the initial drafts of this paper.

Finally, I am indebted to the Institute for Research on Teaching for their financial support during the past year while this research was conducted.

ii

TABLE OF CONTENTS

		Page
LIST OF 1	TABLES	v
LIST OF F	FIGURES	vii
Chapter		
1. 1	INTRODUCTION	1
	Statement of the Problem	12 12
2. 1	THE ANALYTIC STRATEGIES	14
	Gains in Standard Scores	21 23 25 31 35
3. F	REVIEW OF RELEVANT LITERATURE	36
	Data Analysis for Quasi-Experiments Raw Gain Scores Gains in Standard Scores Estimate True Score Analysis of Covariance Analysis of Covariance With Multiple Fallible Covariates Summary	36 39 40 43 49 51
4. E	EFFECTS ESTIMATED AND THEIR STANDARD ERRORS	52
	Estimation of Treatment Effects Estimation With Gains in Standard Scores Estimation With True Score Analysis of Covariance Estimation With Adjusted Gain Scores Estimation With Multiple Fallible Covariates Precision Standard Error For an Index of Response The Standard Error For Kenny's Procedure	55 56 58 61 64 68 69 72
	The Standard Error For Porter's Procedure	77

Chapter

	The Standard Error For Adjusted Gain Scores
5.	DISCUSSION
	A Data Example
Ap pendi	x
Α.	THE VARIANCE OF THE PRODUCT OF TWO RANDOM VARIABLES 116
Β.	THE COVARIANCE OF A RANDOM VARIABLE AND THE PRODUCT OF TWO RANDOM VARIABLES
С.	THE COVARIANCE OF TWO ADJUSTED VARIABLES
D.	THE EXPECTED VALUE OF THE RATIO OF TWO NON-INDEPENDENT SAMPLE STANDARD DEVIATIONS SQUARED
Ε.	THE EXPECTED VALUE OF THE RATIO OF TWO NON-INDEPENDENT SAMPLE STANDARD DEVIATIONS
F.	THE EXPECTED VALUE OF A SAMPLE REGRESSION COEFFICIENT 124
G.	THE EXPECTED VALUE OF A SAMPLE REGRESSION COEFFICIENT SQUARED
Н.	THE VARIANCE OF THE ADJUSTED GAIN SCORE VARIABLE 133
BIBLIOG	RAPHY

LIST OF TABLES

Table		Page
1.	A Hypothetical Data Example of the Fan Spread Hypothesis With the Linear Model of Within-Group Growth	53
2.	A Hypothetical Data Example of the Fan Spread Hypothesis With the Non-Linear Model of Within-Group Growth	54
3.	The Expected Value of the Ratio of Two Non-Independent Sample Standard Deviations When the Population Standard Deviations Are Equal	75
4.	The Variance of the Ratio of Two Non-Independent Sample Standard Deviations When the Population Standard Deviations Are Equal	76
5.	The Variance of a Sample Regression Coefficient When the Population Standard Deviations of the Two Variables Are Equal	81
6.	The Variance of the Adjustment Coefficient Suggested in the True Score Analysis of Covariance Procedure	83
7.	The Standard Errors Associated With Gains in Standard Scores, True Score Analysis of Covariance, and Adjusted Gain Scores	86
8.	Coefficients for the Second, Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When $\rho = .9$ for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal	88
9.	Coefficients for the Second, Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When $\rho = .8$ for the Manifest Variables and the Population Variance of the X and X Variables Are Equal	80
	VI UNG A WING I TWI IWVIGD AND EQUAL I I I I I I I I I I I I I I I I	03

Ta

10.	Coefficients for the Second, Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When ρ = .7 for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal	90
11.	Coefficients for the Second, Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When $\rho = .6$ for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal	91
12.	Coefficients for the Second, Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When $\rho = .5$ for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal	92
13.	Coefficients for the Second, Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When ρ = .4 for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal	93
14.	Group Means in the Metric of Grade Equivalent Scores and Standard Deviations on the Paragraph Meaning Subtest of the Stanford Achievement Test Battery	103
15.	Results of the Data Analyses Using the Gains in Standard Score Strategy, True Score Analysis of Covariance, Adjusted Gain Scores With the Analysis of Variance Model, Adjusted Gain Scores With the Derived Standard Error, Traditional Analysis of Covariance and Analysis of Variance With Residualized Gain Scores	105
16.	The Effect Estimated by the Gains in Standard Scores, True Score Analysis of Covariance and Adjusted Gain Scores, and Traditional Analysis of Covariance	106

Page

LIST OF FIGURES

Figure		Page
1.	The selection by maturation interaction: increasing mean differences in achievement between comparison groups across time	5
2.	The fan spread hypothesis: increasing mean difference in achievement between comparison groups with a proportional increase in the within-group variability across time	6
3.	The fan spread hypothesis with the linear model of within-group growth	8
4.	The fan spread hypothesis with a non-linear model of within-group growth	8
5.	Underadjusting for group differences with the raw gain score procedure in a situation conforming to the differential growth rate theory	17
6.	Differential growth rates considered over three points in time	26
7.	Group achievement means plotted across three points in time	103

C t a r d d 19 Li in Me in ch in The lat CON one inv€

CHAPTER 1

INTRODUCTION

Individuals are dynamic; attitudes, perceptions, and knowledge continually change as a result of interactions among individuals and the communications media. The understanding of this dynamic nature and at times the modification of it, is a major concern of social science research. Interest in naturally changing entities has raised, however, difficult problems in measurement and analysis. Although considerable discussion has been devoted to the topic of measuring change (McNemar, 1958; Lord, 1956, 1958, 1963; Bereiter, 1963; Cronbach and Furby, 1970; Linn & Slinde, 1977), the problem is far from resolved.

The problems related to measuring change exist to varying degrees in all research designs. These issues are less troublesome in experimental studies where the investigator manipulates the variables of interest and the effects on other variables are observed. Measuring change is more difficult in quasi-experimental studies where the investigator lacks the freedom to manipulate the variables of concern. The present study focuses on the issues of change associated with the latter design. Specifically, this study considered the non-equivalent control group design (Campbell & Stanley, 1963) where the results of one or two pretests are available prior to the period under investigation.

Another popular research design frequently adopted by social scientists is based on natural variation. In these studies the investigator identifies a group of individuals and observes them in their "natural environment" on a set of variables of interest. The relationship among these variables is then based on correlational techniques. Measuring change in these studies is more difficult than in quasi-experiments and is not considered here. The finding of the present study may, however, be indirectly relevant to studies based on natural variation.

The evaluation of an educational program frequently results in a quasi-experiment of the type considered here. Compensatory education programs such as Head Start and Follow Through, in particular, have been the focal point of much discussion on growth assessment in quasiexperimental studies (Campbell & Erlebacher, 1970). In these programs students are not randomly assigned to treatments, but rather those in greatest need are given the treatment--that is, additional assistance. Treatment effects are estimated by comparing the academic achievement of those students receiving the additional assistance with a group of students who did not receive the additional assistance. Evaluation in quasi-experimental settings raises a number of issues and problems not generally encountered in true-experimental designs. Campbell and Boruch (1975) discuss in detail several concerns which may arise in the analysis and interpretation of quasi-experimental data. The overriding theme of their work revolves around the issue of bias in the estimation of treatment effects. An estimate of the treatment

effect is biased if the estimate indicates that the effect of the program was either positive or negative when there were no true treatment effects or vice versa. Although there are several factors which contribute to a biased estimate, the entire problem originates from the fact that without randomization there are likely to be substantial differences between the individuals in their initial status on the outcomes to be assessed. Estimating the treatment effects solely on the differences in post-treatment test scores may be biased since, depending on the direction of the initial differences, a program may appear effective or harmful even though there were not true treatment effects. Several strategies have been suggested which take initial differences into consideration when estimating a program's effectiveness. Campbell and Boruch (1975) argue, however, that these adjustment procedures frequently fall short of eliminating all of the bias that can result from initial differences. The magnitude of the bias is related to two issues: (1) specifying the appropriate variables on which the adjustment is made, and (2) identifying the appropriate model to use the variables in predicting change. These are difficult problems to solve. Specifying the appropriate variables means that the investigator can identify those variables which are predictive of all confounding variables that affect the dependent variable. Knowledge of those variables on which the assignment of individuals to groups was based would provide one possible solution. Another solution is the random assignment of individuals to groups. This solution is not possible for studies being considered here.

The present study does not pursue this aspect of the specification problem.

The second issue, that of specifying the appropriate model is a major concern here. This aspect of the specification problem includes the question of measuring the adjustment variables reliably. The unreliability issue has been considered in detail and several solutions have been suggested (Lord, 1960; Porter, 1967; De Gracie, 1968; Stroud, 1972). Specifying the appropriate analytic model is dependent on how individuals change over time. The issue of growth models has recently been considered explicitly by several researchers (Campbell, 1971; Kenny, 1975; Bryk & Weisberg, 1977). Campbell in particular has been concerned with the relationship between growth rates and estimates of treatment effects. Although his interest has centered on the evaluation of compensatory education programs, his work applies equally to other quasi-experimental investigations.

The issue which Campbell has raised involves the implications of initial differences on the outcome dimension for the prospect of differential growth rates. His reasoning is based on the belief that groups which differ in their initial average performance also differ in their rate of development on the outcome dimension. Thus, in the evaluation of compensatory education programs, a control group may have a higher average pretest score because as a group these individuals have grown quicker than the group who will receive the treatment, and this development is likely to continue without program intervention. Pictorially this selection by maturation interaction can be presented

as in Figure 1, where the lines represent group average performance over time.



Figure 1. The selection by maturation interaction: increasing mean differences in achievement between comparison groups across time.

Campbell further develops this idea of differential growth rates into a theory which he has labeled the "fan spread hypothesis." It states that along with the increasing mean difference between the compared groups, a proportional increase in the variance within the groups occurs. Figure 1 can be modified to reflect the changing variance as in Figure 2. The labels for treatment-control are arbitrary.



Figure 2. The fan spread hypothesis: increasing mean difference in achievement between comparison groups with a proportional increase in the within-group variability across time.

The dashed lines represent the increasing range of achievement scores within the treatment and control groups over time. This relationship between the increasing mean difference and the within-group variance can be represented in the following formula:

$$\frac{\mu_{xpt} - \mu_{xct}}{\sigma_t} = K$$

where:

- $^{\mu}xpt$, $^{\mu}xct$ are the population means on measure (X) for the program and control groups, respectively, at time t;
 - σ_t is the pooled within-group standard deviation of the outcome measure at time t; and
 - K is a constant.

Thus the difference between group means relative to the pooled within-group standard deviation remains constant over time. It might be noted here that parallel growth patterns between groups may also conform to Campbell's fan spread model if the within group variance remains constant across time. Evidence supporting the fan spread model of growth has been provided by both cross-sectional and longitudinal studies (Osborne, 1966; Baugham & Dahlstron, 1968; Fennessy, 1974). Although this is a relatively small sample of studies on which to base a theory, Campbell is confident that additional findings will support the model (Campbell & Boruch, 1975). Kenny (1975) has argued for the reasonableness of the theory and has provided additional data conforming to the fan spread hypothesis. This model of growth suggested by Campbell and supported by Kenny presents a special case of a more general issue involving linear-growth models. Bryk and Weisberg (1977) have gone beyond Campbell's fan spread hypothesis and have considered the problem of differential linear-growth patterns. By varying the initial starting points of growth and the average rate of group growth, several linear models were considered.

The discussions of growth models presented by Campbell, Kenny, Bryk, and Weisberg have concentrated on differential growth rates between comparison groups and have ignored the issue of differential growth rates within groups. The question of growth rates within groups can be conceptualized in at least two ways as presented in Figures 3 and 4.



Figure 3. The fan spread hypothesis with the linear model of within-group growth.



Figure 4. The fan spread hypothesis with a non-linear model of within-group growth.

In each diagram the solid line represents the average growth rate for the group, while the dotted lines represent individual rates of growth. Figure 3 presents within-group growth rates that are generally associated with the fan spread hypothesis. It conceptualizes withingroup growth as having a common starting point and different linear rates of individual growth across time. Thus, in any two subsequent points in time, individuals maintain their relative position within the group. Figure 4, on the other hand, represents the situation in which the group's mean growth is linear but individual growth is not linear. Under this model an individual's growth rate may vary over time, i.e., growth may occur in spurts, but group growth may be constant. Both of these models can result in data conforming to Campbell's fan spread hypothesis. The implications these models of within-group growth have for data analysis and estimation of treatment effects are substantially different.

Given that the fan spread model represents a valid conceptualization of how individuals and groups change over time in quasiexperimental studies, Campbell (1971) has argued that current analytic strategies are inadequate in adjusting for the differential nature of growth. As a result, recent efforts to evaluate compensatory education programs may be misleading since the differential growth patterns have not been considered. The conclusion that these programs have been ineffective or even harmful may be a statistical artifact rather than actuality.

In response to Campbell's argument that current analytic strategies inadequately adjust for the fan spread model, several researchers have proposed new or modified techniques to resolve the differential growth problem. Kenny (1975) has argued that given the fan spread model, an appropriate analytic strategy is what he calls standardized gain scores (also referred to as gains in standard scores). The fan spread hypothesis suggests increasing variability within groups across time. Kenny's approach counters this increasing variability by standardizing the pretest and posttest scores using the pooled withingroup standard deviation at time 1 and time 2, respectively. The difference between the standardized scores is then computed and used as the dependent measure with the analysis of variance model.

Another solution to the fan spread model was proposed by Porter and Chibucos (1974). They suggested that the analysis of covariance model was appropriate for the differential growth rate situation if the covariate was perfectly reliable. Given that the covariate was fallible, then analysis of covariance with the estimated true score of the covariate, would adequately adjust for the fan spread model. Estimated true score analysis of covariance was originally developed by Porter (1967) as a solution to the single fallible covariate problem.

Still another solution which might be considered to adjust for the fan spread effect is the use of gain scores adjusted for differential group growth. The raw-gain-score strategy assumes that groups are changing at relatively equal rates, and that the only difference between the groups other than that caused by treatments, is the

initial status at the point of intervention. The fan spread model allows that not only do the groups differ in their pre-treatment performance levels, but also that the groups are changing at different rates. Therefore simple gain scores could be inappropriate in light of the fan spread model. If, on the other hand, the gain scores themselves were adjusted for the differences between the groups' growth rates, an appropriate estimate of the treatment effect might be obtained. An adjustment of the type just described could be made if data from two pretests over time rather than one pretest were available. Given that two pretests were available for each group, then each group's growth rate could be estimated by the difference between the group's mean performance on the first and second pretests.

Given that multiple pretest data are available, a fourth procedure which might be considered to analyze data conforming to the fan spread hypothesis is the analysis of covariance model with multiple covariates. This technique might be used by researchers in the field where the tendency is to use all available information on a group of subjects in the hope of increasing precision and adjusting for all initial differences. Furthermore, in light of the earlier discussion on analysis of covariance, multiple covariates may also adjust for the differential growth rate problem if corrected for their unreliabilities. Recently, Keesling and Wiley (1976) have suggested a procedure which they argue solves the multiple fallible covariate problem. Their procedure may therefore also provide an appropriate solution to the fan_ spread problem.

Statement of the Problem

A great number of educational research efforts are based on designs that are quasi-experimental. As a result, researchers in the field frequently encounter difficult problems in measuring change and estimating treatment effects. Campbell has argued that a difficulty which has not always been explicitly recognized in quasi-experimental studies is a question of differential growth rates. In particular, the evaluation of compensatory education programs, where the children with the slowest academic growth receive the treatment, may be especially vulnerable to this problem. Campbell has suggested that traditional analyses strategies fail to take the differential growth patterns into consideration in estimating treatment effects: thus traditional strategies result in biased estimates of program effectiveness. These programs have therefore appeared less beneficial than was actually the case. In response to Campbell's arguments, several analytic strategies have been suggested which may provide the appropriate adjustment under the fan spread condition.

Purpose of the Study

The purpose of the study was to compare four procedures in terms of their appropriateness as strategies for data analysis in quasiexperimental studies given that individuals and groups may grow differentially. The four strategies considered were these: (1) gains in standard scores, (2) single covariable analysis of covariance with estimated true scores, (3) gain scores adjusted for differential growth rates, and (4) multiple fallible covariable analysis of

covariance. Specifically, two types of individual growth were considered: first, the situation in which the correlation between the pre-intervention and post-intervention measures was unity, $\rho = 1$, except for measurement errors; and second, the situation in which the relationship between the two measures was not perfect, $\rho \neq 1$, regardless of measurement errors. This second case results when individuals begin to grow at different points in time and grow at different rates, or when individuals grow academically in spurts.

The appropriateness of the four strategies was based on two considerations: first, the effect estimated by each technique, and second, the precision with which each effect was estimated. Furthermore, a discussion of the effects and implications of violating either of two assumptions was considered. They were the homogeneity of regression assumption and the assumption that individuals and groups grow in a linear fashion.

In conducting this study two approaches were used. First, the four analysis strategies were considered analytically to determine whether the procedures estimated the effect of interest. Standard errors were then derived for those strategies estimating the appropriate effect to identify those procedures which offered the greatest precision. The second approach in comparing the strategies was to analyze a set of data with the procedures that estimated the appropriate effect. The conclusions drawn by each strategy were then compared as were their respective error terms.

CHAPTER 2

THE ANALYTIC STRATEGIES

The previous chapter has raised explicitly the question of differential growth rates in quasi-experimental studies. While some detail was given to the presentation of this problem, only general statements as to possible solutions to the fan spread model were provided. In this chapter the analytic strategies suggested are considered in detail. Before studying these solutions, however, a brief discussion is presented concerning the inadequacies of raw gain scores and analysis of covariance without correction for an unreliable covariate in quasi-experiments.

In experimental studies where individuals are randomly assigned to a treatment or a control group, program effectiveness can be estimated by the difference between the group means on some measure of interest following the treatment:

$$\alpha = \mu_{yp} - \mu_{yc}$$

Where

 α is the estimate of the treatment effectiveness; μ_{yp}, μ_{yc} are the population means on the post-treatment measure (Y) for the program and control groups, respectively. Since individuals were randomly assigned to the two groups, it can be assumed that prior to the implementation of the treatment, the groups differ only by chance factors on the outcome dimension and their average growth rates were in the long run the same. Furthermore, if the treatment has no true effect, the groups would remain equivalent after the period of program experience and the above estimate would equal zero. For the situation described, the analysis of variance model provides an appropriate analytic strategy to compare the group means in that it estimates the effect of interest. The dependent measure using this strategy is simply the performance on a posttreatment measure on a dimension of interest.

With quasi-experimental studies initial differences are likely on the outcome dimension even before the treatment is implemented. If the above procedure was used to estimate the treatment effect and the program had no true effect, then the initial difference would have been treated as a program effect and erroneous conclusions would have been drawn. Thus in quasi-experimental studies some type of an adjustment is necessary to take into consideration those initial differences which may influence the final outcome measure.

Two strategies frequently chosen by researchers in the field for data analysis in quasi-experimental studies have been the analysis of variance model with raw gain scores as the dependent measure and analysis of covariance. Though the former strategy has often been criticized (Cronbach & Furby, 1970; Campbell & Erlebacher, 1970), it still remains a popular approach for researchers in the field

(Richards, 1975). Raw gain scores are calculated by taking the difference of the post-treatment and pre-treatment scores (post-pre) and creating a new variable. For example, if Y represents the post-treatment score on some variable of interest and X represents the pre-treatment score, a new variable W is created by the simple difference between the two scores; i.e., W = Y - X. Since the new variable is created by taking the difference between scores, this procedure logically requires that the same or equivalent form of the measure be administered at the two points in time. The analysis of variance model is then used with W as the dependent measure. The program effectiveness when the raw gain score strategy is used can be written as:

$$\alpha_{GS} = \mu_{yp} - \mu_{yc} - (\mu_{xp} - \mu_{xc}).$$

Where:

α _{GS}	is the treatment difference estimated by using the gain score strategy;
^µ yp' ^µ yc	are the population means on the post-treatment measure (Y) for the program and control groups, respectively; and
^µ xp' ^µ xc	are the population means on the pre-treatment measure (X) for the program and control groups, respectively.

The adjustment which is made using this technique in a situation conforming to a fan spread model when there are no true treatment effects is presented in Figure 5.



Figure 5. Underadjusting for group differences with the raw gain score procedure in a situation conforming to the differential growth rate theory.

The solid lines represent actual mean growth of individuals in the two groups and the dashed line represents the adjusted growth pattern for G_2 . The difference at t_2 between the solid line for G_1 , and the dashed line for G_2 represents the bias remaining after adjustment in estimating the treatment effect. It is clear that the gain score strategy requires the assumption of equal growth rates between comparison groups in light of no treatment effects. Violation of this assumption results in an underadjustment in the situation depicted and therefore is a biased estimate of the program effect.

A second popular analytic strategy frequently adopted by researchers in the field is the analysis of covariance model. This procedure is similar to the raw gain score strategy in the nature of the adjustment made, but it is not restricted to the use of the same or parallel form of the outcome measure in order to make the adjustment. Rather than subtracting the pretest score from the posttest score, the analysis of covariance model subtracts only a portion of the difference of pretest from the posttest. That portion is equal to the pooled within-group linear regression slope of the line predicting the outcome measure from the adjustment variable. To facilitate the understanding of the effect estimated, the dependent variable can conceptually be thought of as $W = Y - b_{y \cdot X} X$ for the analysis of variance model. Although the adjusted variable is not actually computed, it can be thought of as such to facilitate comparisons across other similar analytic strategies. The treatment difference can then be written as:

$$\alpha_{AC} = \mu_{yp} - \mu_{yc} - \beta_{y \cdot x} (\mu_{xp} - \mu_{xc}).$$
 (1)

Where:

αAC	is the estimate of the treatment difference using the analysis of covariance strategy;
^β y∙x	is the pooled within-group linear regression slope of Y on X; and

 $\mu_{yp}, \mu_{yc}, \mu_{xp}, \mu_{xc}$ are as defined previously.

Campbell and Erlebacher (1970) and Kenny (1975) have argued that in situations conforming to the fan spread hypothesis this strategy also underadjusts for initial differences. As a result, the strategy provides a biased estimate of the treatment effect. Their discussion, however, focused on analysis of covariance with a fallible covariate. Furthermore, the distinction between the two models of within-group growth, that was proposed in Chapter 1 was not made in either presentation. The importance of this distinction is examined in Chapter 4. Bryk and Weisberg (1977), on the other hand, have suggested that the analysis of covariance with a reliable covariate was an appropriate strategy for the fan spread condition. Their conclusion was based on the assumption of linear growth for individuals within treatment groups. The same conclusion was drawn by Porter and Chibucos (1974).

Analysis of covariance has also been criticized as inappropriate in situations conforming to the fan spread model because the assumption of homogeneity of regression slopes between treatment groups may be violated. Bryk and Weisberg (1977) and Campbell and Boruch (1975) have suggested that differential growth rates imply a differential relationship between the covariate and the dependent measure for the comparison groups. These authors, however, may have confused the regression of achievement on time as presented in Figure 1 with the regressions are not the same and a violation of the homogeneity of regression slopes does not have to occur for the latter case. The regression slope for each comparison group is defined as:

$$b_{y \cdot x} = r_{xy} \frac{s_y}{s_x}$$
.

Where

^b y∙x	is the regression slope;
r _{xy}	is the correlation of the pretest and posttest measures; and
s _y ,s _x	is the standard deviation of the posttest and pretest, respectively.

The fan spread model suggests only an increasing variance across time such that $S_y > S_x$. This increase in variance occurs in both groups. Nothing in the differential growth rate situation suggests that the correlation between the two measures should differ from one group to the other. It does not follow then that a violation of the homogeneity of regression slopes assumption is likely. Chapter 4 presents two hypothetical examples conforming to the fan spread model where the assumption is not violated.

Finally, the analysis of covariance strategy has been criticized as inappropriate in quasi-experiments since the slope, reflecting growth over time, is underestimated due to measurement errors in the covariate. This is a legitimate criticism which has received considerable attention as noted earlier. Several solutions to this problem have been suggested. Later in this chapter one of those suggested solutions will be considered in detail.

Thus several methodologists have argued that data analysis for a study based on a quasi-experimental design may be extremely troublesome as a result of differences in growth rates between comparison groups. The evaluation of compensatory education programs, in particular, has been singled out as being especially vulnerable to this problem. Traditional strategies for estimating treatment effects like gain score analysis of variance and analysis of covariance have been criticized as inappropriate since they can underadjust for initial as well as growth related differences between groups. Figure 5 presented a hypothetical example of the bias remaining after using the gain score strategy. It will be shown that a similar underadjustment could arise with the analysis of covariance strategy depending on the nature of growth and the value of $\beta_{v \cdot x}$ in equation 1.

Gains in Standard Scores

There have been several proposals as to how data might be analyzed in light of the fan spread hypothesis. One approach to the problem suggested by Campbell (1971) and later by Kenny (1975) was the use of gains in standardized scores. This approach counters the increasing variability within the treatment and control groups across time by standardizing the pretest and posttest scores separately ("given unit variance and a mean of zero," Kenny, 1975, p. 347). To standardize the scores obtained at each test administration, Kenny suggested the following procedure: from each score, subtract the grand mean (across groups) and divide this difference by the pooled within-group standard deviation. (Since subtracting a constant has no effect on further analysis, it is not considered in the subsequent discussion of the strategy. Rather, standardization is achieved by simply dividing each score by the pooled within-group standard deviation for the pretest and posttest data considered separately.) Conceptually, this results in a new dependent variable W, created by subtracting from the posttest

score the product of the ratio of the pooled within-group standard deviation (posttest to pretest) and the pretest score, i.e., $W = Y - \frac{S_y}{S_x}X$. The variable W is then taken as the dependent measure in the analysis of variance model. The treatment difference can be presented as:

$$\alpha_{GSS} = \mu_{yp} - \mu_{yc} - \frac{\sigma_y}{\sigma_x} (\mu_{xp} - \mu_{xc}).$$

Where:

- α_{GSS} is the estimate of the treatment difference estimated by the gains in standard score strategy;
- σ_y, σ_x are the pooled within-group standard deviations of the pre-treatment and post-treatment measures, respectively; and

 $\mu_{yp}, \mu_{yc}, \mu_{xp}, \mu_{xc}$ are as defined previously.

The above statement of the treatment difference is only an approximation since $E(S) \neq \sigma$. The use of this strategy, like that of raw gain scores, logically requires that the pre-treatment measure be identical to or a parallel form of the post-treatment measure.

In defending the use of this technique, Kenny (1975) presented several examples where standardized gain scores provided an unbiased estimate of program effects while raw gain scores and analysis of covariance were shown to be biased. He concluded that in certain situations (i.e., where individuals were assigned to a program based on sociological or demographic variables) gains in standardized scores provide the best analytic strategy. Bryk and Weisberg (1977) provided further evidence showing that this strategy is appropriate in situations conforming to the fan spread hypothesis.
Analysis of Covariance With Estimated True Scores

Another approach which has been suggested as appropriate in a situation conforming to the fan spread hypothesis is the use of estimated true scores in the analysis of covariance. This procedure was originally developed (Porter, 1967) to eliminate the bias introduced by measurement errors when analysis of covariance is used to estimate treatment effects in quasi-experimental studies. Porter and Chibucos (1974) showed that this procedure corrects for the fan spread theory when the pre-treatment measure predicts the post-treatment outcome perfectly except for errors of measurement. To use the procedure suggested by Porter, the estimated true scores of the covariate must be computed. This can be achieved with the following formula:

$$\hat{T} = \overline{X} + \rho_{XX} (X - \overline{X}).$$

Where:

T is the estimated true score of the covariate; \overline{X} is the group mean on the covariate; X is the covariate's observed score; and ρ_{xx} is the reliability of the covariante.

This approach requires knowledge of the covariate's reliability. The question as to which reliability coefficient should be used in the above formula has not been completely answered, but Porter and Chibucos (1974) have suggested that if possible a test-retest

reliability coefficient over a relatively short period of time should be the first choice. Having estimated the covariate's true score, this variable is then used as the covariate in the analysis of covariance model to estimate group differences.

Using Porter's procedure the program effect can be written as:

$$\alpha_{TS} = \mu_{yp} - \mu_{yc} - \frac{\beta_{y \cdot x}}{\rho_{xx}} (\mu_{xp} - \mu_{xc}).$$

Where:

 ${}^\alpha\text{TS}$ is the estimate of the treatment difference computed by the true score analysis of covariance strategy; and

^μyp^{,μ}yc^{,μ}xp^{,μ}xc[,] $β_{y,x}$ and $ρ_{xx}$ are as defined previously.

The similarity of this approach to that of gains in standard scores presented earlier is clearly shown with the following substitution:

$$\beta_{y*x} = \rho_{xy} \frac{\sigma_y}{\sigma_x}$$

The estimate of the treatment difference can now be written as the following:

$$\alpha_{\text{TS}} = \mu_{yp} - \mu_{yc} - \frac{\rho_{xy}}{\rho_{xx}} \frac{\sigma_{y}}{\sigma_{x}} (\mu_{xp} - \mu_{xc}).$$

Given individuals conform to the fan spread, pre-treatment scores should predict post-treatment scores perfectly except for measurement errors. Thus the ratio of the correlation between measures and the pretest reliability is equal to unity, $\frac{\rho_{XY}}{\rho_{XX}} = 1$ if $\rho_{XX} = \rho_{XY}$. The estimate of the program effect provided by true score analysis of covariance and gains in standard scores is the same for fan spread data conforming to the first model of within group growth. This similarity is only true for the linear growth model for individuals within comparison groups. When individuals within groups are growing non-linearly, the ratio of the correlation between measures and the reliability coefficient of the pretest does not equal unity. The effect estimated by the gains in standard scores and analysis of covariance with estimated true scores is therefore different. The two procedures also differ in that the gains in standard scores approach assumes that the correct ratio of the standard deviations is known for the population, while estimated true score analysis of covariance estimates the parameter on the sample data.

Adjusted Gains Scores

The inadequacies of the raw gain score strategy in situations conforming to the fan spread hypothesis were discussed in some detail earlier. It was shown that for the general fan spread model of growth, gain scores were inappropriate since they adjust only for initial differences at the point of intervention and not for differences in the rate of growth as the model predicts. These latter differences, if uncontrolled, would result in a biased estimate of program effectiveness. The gain score strategy may still be appropriate, however, if modified to reflect not only initial differences but also growth rate differences. Such a modification could be made if additional additional data collected at some time prior to the point of intervention were available. This modified gain score approach then could provide a third alternative solution to the fan spread hypothesis.

To facilitate a discussion on the development of the modified gain score procedure, Figure 6 presents in greater detail differential achievement growth over time for a hypothetical program and control group without a treatment effect.



Figure 6. Differential growth rates considered over three points in time.

The horizontal axis T denotes time while the vertical axis W represents achievement. On the time dimension three points are identified: t_1 , t_2 , and t_3 . The vertical dotted line at t_2 indicates the point of intervention while the dotted lines at t_1 and t_3 represent points in time prior to and at termination of the intervention, respectively. The solid lines represent the linear regression of achievement on time for the populations of program and control groups. In this figure the control group is shown to have a higher achievement rate (growing faster) than the program group. The points at which these growth curves intersect the dotted vertical lines represent the average achievement level on the measure administered at time t. For example, (t_2,μ_{xp}) represents the population mean on measure X for the program group at the time of intervention. These solid lines can be defined in regression equations and used to predict the average group performance at any point in time. If, for example, group performance at t_3 was of interest, the following equations may be used:

 $\mu_{yp} = a_p + b_p(t_3 - t_2)$

and

$$\mu_{yc} = a_c + b_c (t_3 - t_2).$$

where:

- μ_{yp}, μ_{yc} are the population mean performance on measure (Y) at time t₃ for the treatment and control groups, respectively;
 - a_p, a_c are the intercept constants of the growth curves for the treatment and control groups, respectively, at t_2 ;

- b_p,b_c are the slopes (rate of growth per unit time) of the regression line predicting achievement from time for the program and control groups, respectively; and
- $t_3 t_2$ is the period of intervention.

The difference in average performance of the program and control groups at the termination of the intervention can be determined as

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - (a_p - a_c) - [(b_p - b_c)(t_3 - t_2)].$$

When the intervention has no effect, the equation is as follows:

$$0 = \mu_{yp} - \mu_{yc} - (a_p - a_c) - [(b_p - b_c)(t_3 - t_2)].$$

Since the intercepts a_p and a_c of the growth curves are the initial achievement levels prior to the point of intervention, then,

$$a_p - a_c = \mu_{xp} - \mu_{xc}$$

the difference in the mean pretest scores of the two groups. With this substitution the expression becomes the following:

$$\mu_{yp} - \mu_{yc} - (\mu_{xp} - \mu_{xc}) - [(b_p - b_c)(t_3 - t_2)].$$
 (2)

The first two terms of the equation are identical to that of raw gain scores that adjust for initial differences in test performance while the second component adjusts for the differential growth rates. If the slopes are equal, that is the rate of growth is the same for both groups, the second component equals 0 and raw gain scores provide the appropriate adjustment procedure. The fan spread model, however, states that the growth rates are not equivalent, and therefore an additional adjustment is needed. The slope of a regression line is defined as the ratio of the change in the vertical axis to the change in the horizontal axis, i.e., $b = \frac{\Delta W}{\Delta T}$. By using the information available before the start of the intervention, the growth rate for each group can be estimated. For the program group, the regression slope can be written as the following:

$$b_p = \frac{\mu_{xp} - \mu_{zp}}{t_2 - t_1}$$

This equation is the ratio of the change in population mean achievement at two points in time prior to the intervention with the period of time between testing. Similarly for the control group the regression slope can be written as the following:

$$b_{c} = \frac{\mu_{xc} - \mu_{zc}}{t_{2} - t_{1}}$$

With these estimates of growth rates, the third term of expression 2 can be written as:

$$\begin{bmatrix} (b_{p} - b_{c})(t_{3} - t_{2}) \end{bmatrix} = \begin{bmatrix} \left(\frac{\mu_{xp} - \mu_{zp}}{t_{2} - t_{1}} - \frac{\mu_{xc} - \mu_{zc}}{t_{2} - t_{1}} \right) (t_{3} - t_{2}) \end{bmatrix}$$
$$= \begin{bmatrix} \left(\frac{(\mu_{xp} - \mu_{zp}) - (\mu_{xc} - \mu_{zc})}{t_{2} - t_{1}} \right) (t_{3} - t_{2}) \end{bmatrix}.$$

If the period of time between the first and second testing equals the period of intervention t_2 to t_3 , the above equation can be simplified as the following:

$$[(b_p - b_c)(t_3 - t_2)] = [(\mu_{xp} - \mu_{zp}) - (\mu_{xc} - \mu_{zc})].$$

Thus, the difference in group mean gains prior to the period of intervention can be an appropriate estimate of the difference in growth rates between program and control groups. The combination of this adjustment for differential growth rates and the adjustment for differences in initial performance levels results in estimating the treatment effects as:

$${}^{\alpha}AGS = {}^{\mu}y_{p} - {}^{\mu}y_{c} - ({}^{\mu}x_{p} - {}^{\mu}x_{c}) - \left[\frac{({}^{\mu}x_{p} - {}^{\mu}z_{p}) - ({}^{\mu}x_{p} - {}^{\mu}z_{c})}{t_{2} - t_{1}} (t_{3} - t_{2})\right].$$

To achieve this type of adjustment requires the creation of a new variable that takes into consideration all three test results. A variable, W, can be defined as the following:

$$W = Y - X - (\overline{W} - \overline{Z}) \frac{t_3 - t_2}{t_2 - t_1}$$

Where:

 X,Y,t_3-t_2,t_2-t_1 are as defined earlier; \overline{X} is the group mean on the pretest administered at t_2 or the point of intervention; and \overline{Z} is the group mean on the pretest administered at t, or some point in time prior to the intervention. If the time between the first two test administrations equals the period of intervention, the new variable is simply a gain score minus the difference of the group's means at two points in time prior to the intervention. This second factor adjusts for the growth rate of the group. Since this technique makes no assumptions as to the rate of growth or the initial starting point, it is appropriate for any situation where growth is linear for groups.

Analysis of Covariance With Multiple Covariates

The strategy presented above required that two assessments be made prior to the start of the intervention. Given the availability of this pre-treatment information, a fourth procedure which has been suggested to analyze data in a quasi-experimental setting conforming to the fan spread model is analysis of covariance with the two pretests as covariates. It was suggested when this technique was proposed that the covariates should be corrected for their unreliability. Keesling and Wiley (1977) have recently suggested a new approach to the problem which may provide a reasonable solution to the question of multiple fallible covariates. Their approach estimates the treatment effects within groups separately and then compare the magnitude of those effects across groups.

The general model for estimating the adjusted posttest mean for a group can be written as:

$$\tau_{G} = \mu_{yG} - \underline{\Gamma}' \underline{\mu}_{xG} - \underline{\beta}' \underline{\mu}_{TG} .$$

Where:

- τ_{C} ~ is the estimate of the effect for group G;
- ${}^{\mu}yG$ is the population mean on the posttest measure Y for group G;
- $\underline{\mu}_{x}, \underline{\mu}_{T}$ are the vector of population means on the fallible and error free covariates, respectively, for group G; and
- $\Gamma'\beta'$ are the row vectors of structural regression coefficients for the fallible and error free covariates, respectively, for the group G.

In the present study the adjustment procedure using fallible covariates is of primary interest, and therefore a discussion of the error free covariates and their structural coefficients will not be presented here.

In the situation under consideration the posttest mean adjusted by the two pretest means under the Keesling-Wiley model can be written as:

$$\tau_{G} = \mu_{yG} - \gamma_{1}\mu_{x} - \gamma_{2}\mu_{z}$$

Where:

 μ_y, μ_x and μ_z are the population means of the three tests as defined earlier; and γ_1, γ_2 are the structural regression coefficients for the two fallible covariates.

The estimates of the structural regression coefficients are defined as the following:

$$\gamma_{1} = (\sigma_{\xi_{2}}^{2}\sigma_{\xi_{1}\eta} - \sigma_{\xi_{1}\xi_{2}}\sigma_{\xi_{2}\eta}) / (\sigma_{\xi_{1}}^{2}\sigma_{\xi_{2}}^{2} - \sigma_{\xi_{1}\xi_{2}}^{2})$$

$$\gamma_{2} = (\sigma_{\xi_{1}}^{2}\sigma_{\xi_{2}\eta} - \sigma_{\xi_{1}\xi_{2}}\sigma_{\xi_{1}\eta}) / (\sigma_{\xi_{1}}^{2}\sigma_{\xi_{2}}^{2} - \sigma_{\xi_{1}\xi_{2}}^{2})$$

Where:

- ξ_1, ξ_2, η represent the true scores of the first covariate Z, the second covariate X, and the posttest measure Y; and
 - σ^2, σ represent the variance and covariance of the subscripted variables, respectively.

In terms of the notation used elsewhere in this study, the structural regression coefficients can be written as the following:

$$\gamma_{1} = (\beta_{T_{y}} \cdot T_{x} - \beta_{T_{z}} \cdot T_{x} \beta_{T_{y}} \cdot T_{z})/(1 - \rho_{T_{x}}^{2} T_{z});$$

$$\gamma_{2} = (\beta_{T_{y}} \cdot T_{z} - \beta_{T_{z}} \cdot T_{x} \beta_{T_{y}} \cdot T_{x})/(1 - \rho_{T_{x}}^{2} T_{z}).$$

(For a discussion of the distinction between structural equation models and regression models, see Goldberger and Duncan, 1973.) It might be noted at this point that the above regression coefficients are similar to those used by Pravalpruk (1974) in one of the two methods he considered for solving the multiple fallible covariate problem. His estimates were:

$$\gamma_{1} = \left(\beta_{y} \cdot T_{x} - \sqrt{\frac{\rho_{zz}}{\rho_{xx}}}\beta_{T_{2}} \cdot T_{x} \beta_{y} \cdot T_{z}\right) / \left(1 - \rho_{T_{x}}^{2} T_{z}\right)$$

and

$$Y_{2} = \left(\beta_{y} \cdot T_{2} - \sqrt{\frac{\rho_{zz}}{\rho_{xx}}} \beta_{T_{z}} \cdot T_{x} \beta_{y} \cdot T_{x}\right) / \left(1 - \rho_{T_{x}}^{2} T_{z}\right).$$

In his study, however, Pravalpruk only considered situations where the reliabilities of the covariates were equal. Furthermore, since $\beta_{y \cdot T_X} = \beta_{T_y \cdot T_x}$ and $\beta_{y \cdot T_z} = \beta_{T_y \cdot T_z}$, the adjustment coefficients considered by Pravalpruk and those considered by Keesling and Wiley are identical. To obtain his regression coefficients, Pravalpruk first computed the estimated true scores for his covariates and then corrected for the attenuated relationship between the two covariates. Concerning this approach, Pravalpruk concluded that while the correct effect was estimated, the appropriate probability of a type I error was only obtained in a two-group design but not in a four-group design. When a four-group design was considered, the test statistic was shown to be too liberal for practical purposes. Thus Pravalpruk concluded that his method did not provide a satisfactory solution to the multiple fallible covariate problem.

In the approach suggested by Keesling and Wiley the true score variance-covariance matrix is estimated using the replicate measures of the covariates. Using the true score relationship the structural regression coefficients are then estimated. The estimation of these parameters is provided in a computer program by Jöreskog and Van Thillo (1972).

To facilitate comparisons across the analytic strategies proposed for solving the fan spread problem the estimate of a treatment effect can be written as the following:

$$\alpha_{MAC} = (\mu_{yp} - \mu_{yc}) - \frac{\rho_{T_x T_y} - \rho_{T_x T_z} \rho_{T_z T_z}}{1 - \rho_{T_x T_z}^2 T_z} \left(\frac{\sigma_{T_y}}{\sigma_{T_x}}\right) (\mu_{xp} - \mu_{xc})$$
$$\frac{\rho_{T_z T_y} - \rho_{T_x T_z} \rho_{T_x T_z}}{1 - \rho_{T_x T_z}^2 T_z T_z} \left(\frac{\sigma_{T_y}}{\sigma_{T_z}}\right) (\mu_{zp} - \mu_{zc})$$

While the above procedure has been demonstrated on a data set, there have not as yet been any investigations considering the properties of the distribution of the test statistic in small samples. Although further study of the Keesling-Wiley procedure is needed before it can be adopted as a competing analytic strategy, it was considered in the present study because it appeared to be a promising technique for the future.

Summary

The four analytic strategies proposed as solutions to the fan spread model have been considered in some detail in this chapter. The following chapter reviews the literature concerning data analysis in quasi-experiments, focusing on discussions directly relevant to the strategies considered in the present chapter.

CHAPTER 3

REVIEW OF RELEVANT LITERATURE

The first chapter began by identifying the measurement of change as a difficult problem in a study based on a quasi-experimental research design. As an example of this type of study, the evaluation of compensatory education programs was suggested. These investigations are characterized by the fact that comparison groups frequently differ in their initial status on the outcome of interest which generally is some measure of achievement. It was then pointed out that some authorities have suggested that differences in initial achievement levels were an indication that the groups were growing academically at different rates. Campbell has labeled this differential growth rate issue the fan spread hypothesis. Given this model, traditional analytic strategies have been questioned. Several procedures were then introduced as potential solutions for differences in growth rates. These solutions were described in detail in Chapter 2. The present chapter reviews the literature as it relates to the proposed solutions.

Data Analysis for Quasi-Experiments

The debate on the analysis of quasi-experiments has ranged from pessimistic arguments stating that appropriate analytic procedures do not currently exist for these studies (Lord, 1967; Crombach & Furby,

1970) to the optimistic view that a cautious interpretation of carefully analyzed data is useful (Elashoff, 1969; Harnquist, 1968; Porter, 1973). Campbell and Erlebacher (1970) reviewed and explicitly demonstrated using simulated data, many of the problems raised when treatment effects are estimated from quasi-experiments. The purpose of their presentation was to illustrate the inappropriateness of gain scores and analysis of covariance for estimating the effectiveness of compensatory education programs. Campbell and Erlebacher observed that pretest data often showed control groups having a higher average performance than those individuals involved in the compensatory program. "That difference no doubt is there because of previously more rapid rate of growth on the part of the Control Group, which would be expected to continue during the period of the experimental treatment" (p. 198). Thus in addition to all of the other problems the literature has identified as being present in quasi-experimental studies, differential growth rates of treatment groups must also be considered in the selection of an analytic strategy. Campbell and Erlebacher suggested that as a result of all of these issues, traditional analyses techniques were inadequate to estimate the effectivenss of compensatory education programs. When these procedures were used, biased estimates resulted, and programs appeared to be ineffective or even harmful.

In discussing the fan spread hypothesis, Campbell and Erlebacher have focused on differences in growth rates between comparison groups and have neglected the question of within group growth. Their argument that analysis of covariance is inappropriate for the fan spread model

is dependent on the nature of the within group growth. This issue is examined in detail in the next chapter. The authors also have confused the problem of differential growth rates with the problem of an unreliable covariate. Given the linear model of within group growth, analysis of covariance provides an appropriate adjustment if the covariate is perfectly reliable. Finally, Campbell and Erlebacher do not point out that parallel growth rate for comparison groups is a special case of the fan spread model when the variance of the measure remains constant across time.

Building on many of the same arguments raised by Campbell and Erlebacher, Campbell and Boruch (1975) identified six ways in which quasi-experimental research designs in evaluations of compensatory education programs underestimate program effectiveness. The issue of differential growth rates among treatment groups was cited as a major contributing factor to a biased estimate of a program effect. Although the authors attempted to clarify the question of differential growth rates, they fell short of addressing those issues ignored in the Campbell and Erlebacher paper. The authors did provide, however, several examples of studies from the literature which supported their contention that cognitive test scores do follow fan spread patterns. As a possible solution to this problem, the authors suggested that "standardizing scores (to mean zero and variance one) at each age level would produce a metric which eliminates differential growth rates" (p. 37). While this procedure may be appropriate, Campbell and Boruch argued that actual growth patterns in achievement were not very well understood and deserved further study.

Raw Gain Scores

Several authorities (Lord, 1956; McNemar, 1958; Cronbach & Furby, 1970; Marks & Martin, 1973) have considered the question of measuring true change with gain scores. Their concern, however, has been with individuals rather than groups. Although related, a distinction should be made between the two topics. The literature has not always done this with the result of conflicting statements and confusion as to the appropriateness of analysis strategies.

The use of gain scores as a strategy to measure group change has been questioned by Cronbach and Furby (1970). These authors pointed out that in true experiments only differences in the post-treatment tests were needed to determine a treatment effect. If a pre-treatment measure was available, then other more powerful techniques such as analysis of covariance provided more powerful estimates of treatment effects than gain scores. In quasi-experimental studies Cronbach and Furby argued that the distribution of true pre-treatment measures was different across the subpopulations being compared. As a result, these authors agreed with Lord (1967) and concluded that no analysis strategy was appropriate to compare treatment effects in quasi-experiments.

Not all authorities share this pessimistic view concerning gain scores as an analytic strategy in quasi-experimental studies. Porter (1973) who has argued that the use of gain scores does not provide the best strategy in true experimental studies, has suggested that under certain assumptions gain scores may provide the best technique for quasi-experimental investigations. The issue of the appropriateness

of the technique is dependent on the reasonableness of the assumption that the relationship between the true pre and true post measures is actually unity ($\beta_{T_y} \cdot T_x = 1$). Porter argued that given the assumption that the treatment effects are additive and that pretest and posttest measure the same variable in a common metric, it can be shown that the relationship between the true parts of the measures does equal unity. He concluded that given these assumptions, the gain score strategy does provide a reasonable approach to data analysis in quasi-experimental studies.

The situation in which Porter has suggested that gain scores would be appropriate is the special case of the fan spread model when the growth patterns are parallel and the variance of the measures remain constant across time. The additivity assumption implies the linear within group growth model.

Gains in Standard Scores

As noted earlier, a solution to the more general fan spread model was proposed by Campbell and Boruch (1975). Their solution was to standardize the observation at each point in time. Although they indicated that this procedure may be appropriate they did not recommend its use, but rather suggested that further study of academic growth patterns is needed. Kenny (1975) took a much stronger position on Campbell's proposed solution and argued that under certain conditions standardized gain scores provided the best analytic strategy for quasiexperiments. Kenny pointed out that in quasi-experiments the procedure assigning groups to a treatment was the determining factor in selecting

the appropriate analysis strategy. For example, if subjects were assigned to the treatment strictly on test scores to be used as the covariate then analysis of covariance without adjustments for the unrealiability of the covariate was the appropriate analysis technique. On the other hand, if the subjects themselves determined which treatment they received, then analysis of covariance corrected for a fallible covariable was the appropriate approach. Finally, if the treatment was assigned to groups based on sociological or demographic characteristics of the group, then the use of standardized gain scores was the appropriate procedure. An example of a study based on sociological or demographic variables is one in which members of a particular social group are eligible for treatment as a matter of legislation. Another example is a study in which treatment is given to members of a particular organization such as a school district. A current example where a compensatory education project was assigned by legislation is the Response to Educational Needs Project in the Anacostia region of Washington, D.C. This program was initiated by the President and Congress specifically for this area in order to increase community involvement and the quality of education in the schools of the region. (This project is currently being evaluated by the National Institute of Education.)

Kenny suggested that the assigning of groups to treatments based on sociological and demographic factors resulted in situations conforming to the fan spread model. That is, treatment and control groups were likely to have different growth rates. Kenny judged the

use of raw gain scores, analysis of covariance, and adjusted analysis of covariance to be inappropriate strategies for these types of studies. Standardized gain scores, however, were appropriate since the technique adjusts for the increasing variability across time. To support his contention as to the appropriateness of the procedure, Kenny cited two examples in an attempt to demonstrate the inappropriateness of raw gain scores and analysis of covariance with or without correction for the fallible covariate while the standardized gain score strategy drew the correct conclusion. He thus concluded that the use of standardized gain scores was an appropriate technique for many quasi-experimental studies.

Kenny's presentation on the appropriateness of the procedure he advocates has several weaknesses. First the label of standardized gain scores is misleading. Gains in standard scores is a more accurate description of the approach. A second weakness is the author fails to distinguish different models of within group growth. As a result his conclusion that adjusted covariate analysis of covariance does not estimate the treatment effect correctly is in error. This issue is addressed in Chapters 4 and 5. A third problem with Kenny's paper is that after showing that gains in standard scores appropriately estimates the treatment effect of interest, the author fails to consider the standard error of the approach. As a result of this oversight, Kenny failed to recognize that the procedure he recommended results in a spuriously low standard error. The standard error issue is considered in detail in Chapters 4 and 5.

Estimate True Score Analysis of Covariance

The second solution to the general fan spread model that was proposed in the first two chapters was the estimated true score analysis of covariance. The technique differs from traditional analysis of covariance in that the estimated true score rather than the observed score on the pre-treatment measure is used as the covariate in the analytic model.

Analysis of covariance with an observed score covariate has been criticized as inappropriate as a result of measurement errors. The problem arises in the calculation of the slope of the regression of the posttest on the pretest used in computing the residuals. The treatment effect which is of interest is based on the latent true variables, but only the observed scores are available. If the regression slope estimated on the observed scores is not the same as that which would be obtained if the latent true scores were available, the regression estimate is said to be biased, and the effect which is estimated is not likely to be the desired one. In the classical analysis of covariance model the covariate is assumed to be fixed rather than random. That is, the levels of the covariates are chosen by the researcher. Berkson (1950) considered using a controlled observation (fixed levels) as the independent variable in estimating the regression line. He concluded the regression slope in this situation to be unbiased. On the other hand, when the independent variable is uncontrolled (a random sample from an existent population), the slope is biased. "The distinction between the two situations

resides not in what the variate represents but how the observations are acquired. If the values are obtained by taking a sample from an existent population, we have the biased situation. If they are obtained by making them as controlled observations, we have the biased situation" (Berkson, 1950, p. 179). In education and psychology, it is rare to find a situation where the covariate is fixed. Rather, for most cases the covariate is a random variable and it is usually measured with error. Thus the slope which is estimated on the observed scores is biased, and the effect which is estimated is not the one of prime interest when treatment groups differ on their mean covariate score. The nature of the bias in using the observed scores rather than the latent true variables can be shown in a number of ways. From a measurement perspective, it is most easily seen through the following set of identities:

$$\beta_{y \cdot x} = \rho_{xy} \frac{\sigma_y}{\sigma_x}$$
(3)

$$\rho_{xy} = \rho_{yT_x} \sqrt{\rho_{xx}}$$
(4)

$$\sigma_{\mathbf{x}} = \sigma_{\mathbf{T}_{\mathbf{x}}} / \sqrt{\rho_{\mathbf{x}\mathbf{x}}}$$
(5)

$$\beta_{y \cdot x} = \rho_{yT_{x}} \sqrt{\rho_{xx}} \sqrt{\rho_{xx}} \frac{\sigma_{y}}{\sigma_{T_{x}}}$$
(6)

$$\beta_{\mathbf{y}\cdot\mathbf{x}} = \rho_{\mathbf{x}\mathbf{x}} \beta_{\mathbf{y}\cdot\mathbf{T}_{\mathbf{x}}}.$$
 (7)

Equation 7 shows that when the observed scores are used to estimate the slope of the regression of the dependent measures on the covariate, it underestimates the desired relationship based on the latent true variables. The observed score regression is biased by a factor equalling the reliability of the covariate. The treatment effect estimated by the observed score analysis of covariance can be written as

$$\alpha_{AC} = \mu_{yp} - \mu_{yc} - \beta_{y \cdot x} (\mu_{xp} - \mu_{xc}).$$

The treatment effect estimated on the latent true variables, on the other hand, can be written as

$$\alpha_{TS} = \mu_{yp} - \mu_{yc} - \frac{\beta_{y \cdot x}}{\rho_{xx}} (\mu_{xp} - \mu_{xc}).$$

Since the latent slope is always larger than the observed score slope, the difference between the two group means on the posttest is adjusted to a lesser extent with the unadjusted regression slope. In the case of evaluating compensatory education programs, analysis of covariance with observed scores underadjusts for initial differences between comparison groups mean on the covariate.

To obtain an unbiased estimate of the regression slope, Lord (1960) suggested a large sample approach which required two independent measures of the covariate. The resulting test statistic, U, was shown to follow the normal distribution and provide an unbiased estimate of treatment effects. This procedure is limited to studies comparing two groups with large samples and two measures on the covariate. Porter (1967) studied the distributional properties of the U statistic when samples were small. His results indicated that a sample size larger than 20 was necessary for the U statistic to converge to the normal distribution when the reliability of the covariate and/or the correlation of the covariate and dependent measure were low.

Another approach to this issue of bias was suggested by Harnquist (1958). He proposed that the problem of an unreliable covariate could be solved simply by multiplying the sum of squares of the covariable by the estimate of the covariable's reliability. Porter (1967), in further developing this idea, suggested that the estimated true score of the variable be used as the covariate. Using this approach, Porter showed that this technique provides an unbiased estimate of the regression slope. Furthermore, in a simulation study using this technique, he showed that the analysis of covariance strategy using true scores followed very closely the theoretical F-distribution.

Stroud (1972) provided still another solution to the problem of an unreliable covariate but his procedure required knowledge of the error variance of the covariate. This approach is also a large sample solution and as yet no small sample distributional investigations have been done with this procedure.

The fan spread model as suggested by Campbell is a special conceptualization of a more general linear growth model. Several models of linear-group growth, including the fan spread, were considered in detail by Bryk and Weisberg (1977) for the non-equivalent control group design. These authors suggested that a simple representation of individual linear growth could be provided by considering

the product of growth rate (π) and time, i.e., the difference between (t_i) the point of assessment and (τ) the initial starting point of growth; growth = $\pi(t_i - \tau)$. Similarly, the average performance of a group at time t_{l} was defined as $\mu_{\pi}(t_{l}$ - $\mu_{\tau})$ where μ_{π} was the average growth rate for individuals within the group and μ_{τ} was the average starting point within the group. By changing the initial starting point for the groups and their rate of growth, Bryk and Weisberg considered five linear growth models. To produce the fan spread condition, a common starting point was assumed for all individuals and the average rate of growth was assumed constant for the group but varied between the groups. Growth within the groups was assumed constant for the individual but varied across individuals. Using this model of growth, the authors compared four analyses strategies on the basis of the effect estimated by each technique. The procedures which were considered were the following: (1) raw gain scores, (2) gains in standard scores, (3) analysis of covariance, and (4) Belson's approach to analysis of covariance. Belson's method, in contrast with traditional analysis of covariance, estimates the regression slope using only the data on the control group. In contrast to Kenny's conclusions, Bryk and Weisberg's analysis indicated that both gains in standard scores and analysis of covariance adequately adjusted for the fan spread model. This conclusion was based on the assumption that the covariate was perfectly reliable. When considering the other models of linear growth that were studied, the authors concluded that these approaches were inadequate in estimating treatment effects. They

therefore suggested that pretest-posttest data were inadequate to adjust for initial differences that frequently accompany the non-equivalent control group design. Bryk and Weisberg further recommended that researchers develop new procedures which might take into consideration more data on growth collected prior to an intervention.

In their study, Bryk and Weisberg made a basic assumption that both groups and individuals within groups grow linearly. As a result they showed that analysis of covariance with a reliable covariate correctly adjusted for the differential growth rates. The assumption that individuals within groups grow linearly, however, is not necessary for the fan spread situation as was pointed out in the first chapter. Without this assumption and the assumption of a perfectly reliable pretest, however, the procedure would not provide the appropriate adjustment. These assumptions were not explicitly stated or discussed in the authors' presentation. Furthermore, Bryk and Weisberg in concluding that both gains in standard scores and analysis of covariance provide the appropriate adjustment, implied that the two procedures are equally appropriate. This is not completely true. While both procedures estimate the same effect, the precision with which the estimate is made differs. The precision of the strategy suggested by Kenny is spuriously high. The issue of precision is considered in detail in the next two chapters.

Analysis of Covariance ith Multiple Fallible Covariates

The fourth analytic strategy which was proposed as a solution to the fan spread model was the analysis of covariance with multiple covariates. This procedure is an extension of the single covariate analytic technique previously discussed. As in the single covariate approach, multiple covariate analysis of covariance has been challenged as inappropriate because of the unreliability of the measures. Unfortunately, the solution to this problem appears more complex than the single fallible covariate case. Pravalpruk (1974) considered two approaches to this problem when two fallible covariates were available. The first method he considered was an extension of Porter's (1967) procedure of using estimated true scores as covariates. In addition, however, he corrected the correlation between the two covariates for attenuation before calculating the beta weights. In the second approach considered, Pravalpruk transformed the second fallible covariate to be independent of the first fallible covariate and then used the estimated true score procedure. The first method was shown to estimate the desired effect, while the second method did not. In simulating the distributions of the F-statistic computed by the two methods, he found that in the case of two-group designs the probability of a type I error rate was satisfactory. In the case of a four-group design, however, the type I error rates were found to be too liberal for practical use. Pravalpruk concluded that the problem of multiple fallible covariate remains unsolved.

Stroud (1974) extended the procedure he suggested to the single fallible covariate problem to include multiple covariates. His solution, however, requires knowledge of the error variances associated with each covariate. As yet no investigations have been conducted studying the distributional properties of his test statistic when small samples are available.

More recently, Keesling and Wiley (1976) have suggested a procedure to adjust for initial differences that may accompany the non-random assignment of experimental units to treatment groups. Their technique is an extension of Lord's (1960) large sample covariance analysis (discussed earlier) to include multiple covariates, some of which may be error-free and others fallible. The two approaches are similar in that the Keesling-Wiley procedure is also a large sample solution; it requires replicate measures of the fallible covariates, and it is useful in the analysis of two-group designs. Unlike Lord's technique, however, the Keesling-Wiley procedure requires replicate measures of the dependent variable as well as the covariates. It does not require that the replicate measures be parallel, nor is it limited to two replicates per variable.

Briefly, the Keesling-Wiley strategy is to take the withincell observed variance-covariance matrix of all replicate measures to estimate the within-cell true score covariance matrix. Using these estimates of true score relationships, the parameters of the structural regression system of true dependent variables on the true covariates can then be computed. These structural regression coefficients are

maximum-likelihood estimates of the parameters computed following the procedure suggested by Jöreskog (1973) and implemented in a computer program by Jöreskog and Van Thillo (1972). The true adjusted dependent variable is then calculated for each treatment group and compared to assess the treatment differences. Further discussion of this approach was presented in the previous chapter.

Summary

The review of the literature has indicated that the measurement of change and the estimation of treatment effects in quasi-experimental studies has been a focal point of a great deal of discussion. In discussing these issues, researchers have shown considerable disagreement as to the appropriateness of various analytic strategies suggested for quasi-experiments. The lack of explicit recognition of differential growth rates as a contributing factor to biased estimates of program effectiveness has contributed to that debate. The literature review has also identified a serious weakness in the previous work in the area of analytic strategies for quasi-experiments. That deficit concerns the almost total neglect of precision as an issue in selecting among competing analytic strategies. The next two chapters attempt to shed some light on the issue of precision as it relates to the four proposed solutions for the fan spread condition.

CHAPTER 4

EFFECTS ESTIMATED AND THEIR STANDARD ERRORS

The four strategies considered in the second chapter have been suggested as possible solutions to the problem of differential growth rates between comparison groups in quasi-experimental studies. The present chapter compares and evaluates these strategies on the basis of two criteria: (1) the appropriateness of the effects estimated, and (2) the precision with which the effects are estimated. While the first criterion must be met before the issue of precision can be sensibly addressed, the second criterion provides a useful basis on which to decide among appropriate competing analytic strategies.

Two conditions of within-group growth are considered in evaluating the effect estimated by each strategy. The first assumes growth to be linear at the individual level across time such that an individual's relative position within the group remains constant over time. This is the "traditional" model of the fan spread condition examined by Campbell and others and represented pictorially in Figure 4 found in Chapter 1. This model of growth suggested that individuals began to grow at some common point in time; and that the rate of growth varied among individuals, but is constant for an individual within a group. As a result, the relationship between test performance across time is perfect except for measurement errors, $\rho = 1$. Table 1 depicts

Table 1

	Individual	Time O Z	Time 1 X	Time 2 Y
Group 1	1	1.8	3	4.2
	2	3.0	5	7.0
	3	4.2	7	9.8
Mean		3.0	5	7.0
Standard deviation		1.2	2	2.8
Group 2		4.8 6.0 7.2	8 10 12	11.2 14.0 16.8
Mean		6.0	10	14.0
Standard deviation		1.2	2	2.8

A Hypothetical Data Example of the Fan Spread Hypothesis With the Linear Model of Within-Group Growth

this model of within-group growth for two hypothetical groups, each consisting of three individuals at three points in time. For this data set, measurement errors are assumed to equal zero. The second condition of within group growth which is considered assumes that individuals within a group do not grow linearly but rather in "spurts." The rate of growth for an individual may therefore vary across time resulting in some variability in the relative position an individual holds within the group. This view of the fan spread model has not previously been considered in the literature. Pictorially it was represented in Figure 5 found in Chapter 1. Following this model of within-group growth, the relationship between test performances across time is no longer perfect even without measurement errors, $\rho \neq 1$. Table 2 presents this model of

	Individual	Time O Z	Time 1 X	Time 2 Y
Group 3	1	2	11	22
	2	3	17	31
	3	4	15	25
	4	5	18	28
	5	6	19	34
Mean		4	16	28
Standard deviation		1.56	3.16	4.74
Group 4	6 7 8 9 10		15 21 19 22 23	28 37 31 34 40
Mean		6	20	34
Standard deviation		1.56	3.16	4.74

A Hypothetical Data Example of the Fan Spread Hypothesis With the Non-Linear Model of Within-Group Growth

Table 2

within-group growth for two hypothetical groups each consisting of five individuals at three points in time. For these data it is assumed that there are no errors of measurement. While the two conditions described above differ in terms of within-group growth rates, both have average group growth that is linear. In considering both the appropriateness of the effect estimated and the question of precision, it has been assumed that the measures available are the same or parallel forms of the same instrument.

Estimation of Treatment Effects

The fan spread model of growth, as discussed earlier, suggests that concomitant with an increase in mean difference between comparison groups is a proportional increase in within-group variability. Furthermore, this relationship between the mean differences and pooled standard deviation remains constant across time. Algebraically this relationship is presented as

$$\frac{\mu_{xp} - \mu_{xc}}{\sigma_{x}} = \frac{\mu_{yp} - \mu_{yc}}{\sigma_{y}}$$

The terms are as defined previously. This representation of the differential growth rate problem indicates that the appropriate adjustment strategy should take the following form:

$$\mu_{yp} - \mu_{yc} - \frac{\sigma_y}{\sigma_x} (\mu_{xp} - \mu_{xc}).$$

An analytic strategy having the above form would provide an unbiased estimate of group differences in situations conforming to the fan spread model of growth. Since the definition of the fan spread hypothesis does not include a reference to the nature of the within-group growth pattern, the above approach is appropriate for both condition 1, $\rho = 1$, and condition 2, $\rho \neq 1$. Assuming that group 1 represents a program group and group 2 the control, the estimate of a non-existent treatment effect for the data from Table 1 depicting condition 1 is as follows:

$$\alpha = (\mu_{yp} - \mu_{yc}) - \frac{\sigma_y}{\sigma_x} (\mu_{xp} - \mu_{xc})$$

$$(7 - 14) - \frac{2.8}{2.0} (5 - 10)$$

$$0 = -7 - 1.4 (-5).$$

Similarly assuming that group 3 represents a program group and group 4 its control, the estimate of the group differences for the data in Table 2 representing condition 2 is as follows:

$$\alpha = (\mu_{yp} - \mu_{yc}) - \frac{\sigma_y}{\sigma_x} (\mu_{xp} - \mu_{xc})$$

$$(28 - 34) - \frac{4.74}{3.16} (16 - 20)$$

$$0 = (-6) - 1.5 (-4).$$

The above examples demonstrate the appropriateness of the adjustment technique suggested by the definition of the fan spread model or alternatively that the data conform to the fan spread model.

Estimation With Gains in Standard Scores

The nature of the hypothesis tested by each of the four analytic strategies is reflected in their respective estimates of group differences presented in Chapter 2. The gains in standard scores approach suggested by Kenny was shown to estimate group differences as

$$\alpha_{GSS} = \mu_{yp} - \mu_{yc} - \frac{\sigma_y}{\sigma_x} (\mu_{xp} - \mu_{xc}).$$

This equation is identical to the adjustment strategy suggested above based on the definition of the fan spread model of growth. In Campbell's definition of the fan spread model, however, it was not clear whether or not the hypothesis was based on manifest or latent variables. If the fan spread hypothesis is defined on the latent true variables then gains in standard scores uses the ratio of the standard deviations on the observed variables, $\frac{\sigma_Y}{\sigma_X}$, when the ratio of the standard deviations of the true variables, $\frac{\sigma_{TY}}{\sigma_{T_X}}$, is desired. The relationship between the variance of the manifest variables to the variance of the latent true variables are shown in the following expressions:

$$\sigma_{\mathbf{X}}^2 = \sigma_{\mathbf{T}_{\mathbf{X}}}^2 \rho_{\mathbf{X}\mathbf{X}}$$

and

$$\sigma_y^2 = \sigma_T^2 \rho_{yy}.$$

The ratio of the standard deviations on the manifest variables in terms of the latent true variables can be written as the following:

$$\frac{\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}} = \frac{\sigma_{\mathbf{T}}}{\sigma_{\mathbf{T}}} \sqrt{\rho_{\mathbf{x}\mathbf{x}}}$$

If the reliability of the pretest equals the reliability of the posttest, $\rho_{XX} = \rho_{yy}$, then the ratio of the observed score standard deviations is appropriate for the latent fan spread model. On the other hand for the latent fan spread model, the ratio of observed standard deviations is an inappropriate adjustment coefficient when the reliability of the measures are not equal. Under the manifest fan spread model, gains in standard scores strategy as proposed by Kenny is appropriate regardless.

While the discussion of the fan spread model indicated that the appropriate adjustment coefficient was the ratio of the population standard deviations for the posttest to the pretest, Kenny's procedure estimates that ratio using the sample standard deviations. The expected value of the ratio of the sample standard deviations, however, does not equal the ratio of the population standard deviation, $E(S_y/S_x) \neq \sigma_y/\sigma_x$, where the samples are small. The effect estimated by Kenny's technique, therefore, is not the desired one when the sample size is small. Further discussion of this point is presented later in the chapter. A final consideration concerning the gains in standard scores approach is the fact that this procedure is not affected by the relationship among the individuals within the comparison groups. For large samples, then, the technique estimates the appropriate effect for both models of within-group growth.

Estimation With True Score Analysis of Covariance

The second solution to the fan spread hypothesis proposed earlier was the analysis of covariance model with estimated true scores as the covariate. This approach estimates the group differences as

$$\alpha_{\text{ACTS}} = \mu_{yp} - \mu_{yc} - \frac{\rho_{xy}}{\rho_{xx}} \frac{\sigma_{y}}{\sigma_{x}} (\mu_{xp} - \mu_{xc}).$$
This statement of the estimate of group differences indicates that this strategy is identical to the adjustment strategy suggested by the fan spread definition except for the ratio of ρ_{XY}/ρ_{XX} . This ratio provides a correction for errors of measurement. Therefore, if the true relationship between the two measures is perfect as proposed by condition 1, the ratio of the correlation to the reliability of the covariate will also equal unity. Thus for the first condition being considered, the analysis of covariance model with estimated true scores provides the appropriate adjustment for the fan spread situation. This result contradicts Kenny's (1975) conclusion that analysis of covariance adjusted for an unreliable covariate does not appropriately adjust for the differential growth rate problem. Since Kenny did not specify the model of within group growth in his study, it must be assumed that his examples illustrated situations where individual growth rates were not linear across time.

Previously a distinction was drawn between the manifest and latent fan spread models. Considering the latent model, the relationship between the adjustment coefficient provided by estimated true score analysis of covariance and the ratio of the latent variable standard deviations can be written as

$$\frac{\rho_{\mathbf{X}\mathbf{y}}}{\rho_{\mathbf{X}\mathbf{x}}}\frac{\sigma_{\mathbf{y}}}{\sigma_{\mathbf{X}}} = \frac{\sigma_{\mathbf{T}\mathbf{y}}}{\sigma_{\mathbf{T}\mathbf{x}}}$$

The above expression is true when the reliability of the pretest and posttest are equal for the linear model of within group growth. If

the reliabilities are not equal then the following set of relationships show that the appropriate adjustment is still provided by the procedure:

$$\frac{\rho_{xy}}{\rho_{xx}} \frac{\sigma_{y}}{\sigma_{x}} = \frac{\rho_{xy}}{\rho_{xx}} \left(\frac{\sigma_{y}}{\sqrt{\rho_{yy}}} \right) / \left(\frac{\sigma_{x}}{\sqrt{\rho_{xx}}} \right)$$
$$= \frac{\rho_{xy}}{\rho_{xx}} \frac{\sigma_{y}}{\sigma_{x}} \frac{\sqrt{\rho_{xx}}}{\sqrt{\rho_{yy}}}$$
$$= \frac{\rho_{xy}}{\sqrt{\rho_{xx}} \rho_{yy}} \frac{\sigma_{y}}{\sigma_{x}}$$

The last expression equals the ratio of the latent standard deviations when the linear model of within group growth is true. If the manifest fan spread model is assumed then the effect estimated by the true score analysis of covariance strategy is appropriate only when the reliabilities of the pretest and posttest are equal.

The data example presented earlier from Table 1, where no errors of measurement were assumed and $\rho = 1$, demonstrated the appropriateness of the true score analysis of covariance technique. The example also demonstrated that the assumption of the homogeneity of regression slopes between comparison groups need not be violated in a fan spread situation. For both groups the relationship between the X and Y measures equaled one and the standard deviation for each variable was equal across the two groups. The regression slopes are therefore equal between comparison groups.

Under condition 2, however, the true relationship between the pretest and posttest does not equal unity even after correcting for

errors of measurement. The data in Table 2 represent this type of situation. In this example the true relationship between the X and Y measures equals .90. The estimate of the adjusted group difference determined by the analysis of covariance model with estimated true scores of the covariate is as follows:

$$\alpha_{\text{ACTS}} = (\mu_{yp} - \mu_{yc}) - \frac{\rho_{xy}}{\rho_{xx}} \frac{\sigma_y}{\sigma_x} (\mu_{xp} - \mu_{xc})$$
$$= (28 - 34) - (.90) \left(\frac{4.74}{3.16}\right) (16 - 20)$$
$$= (-6) - (1.35) (-4)$$
$$= -.60.$$

Thus under condition 2 this strategy underadjusts for initial group differences. The data in Table 2 also demonstrate the assumption of homogeneity of regression slopes does not have to be violated in situations conforming to the fan spread model.

Estimation With Adjusted Gain Scores

The third analytic strategy proposed in Chapter 2 to provide an appropriate adjustment for the differential growth rate problem was the adjusted gain score stratagy. This technique suggested that given the results of two test administrations prior to the introduction of the treatment, the average rate of growth within each group could be estimated and used to adjust the raw gain scores for differences in the growth rates. Estimating group differences with the adjusted gain score strategy was represented as

$$\alpha_{AGS} = (\mu_{yp} - \mu_{yc}) - (\mu_{xp} - \mu_{xc}) - [(\mu_{xp} - \mu_{zp}) - (\mu_{xc} - \mu_{zc})] \frac{t_2 - t_1}{t_1 - t_0}$$

The terms are as previously defined. In situations where the period of time between the first and second pretests, $(t_1 - t_0)$ equals the period of intervention, $(t_2 - t_1)$, the above estimate can be simplified as the following:

$$\alpha_{AGS} = (\mu_{yp} - \mu_{yc}) - 2(\mu_{xp} - \mu_{xc}) + (\mu_{zp} - \mu_{zc}).$$

The utility of the effect estimated is demonstrated for the fan spread data by showing that $\alpha_{AGS} = 0$. Assume that the difference between the two group means on the X variable equals some constant (a), $\mu_{XP} - \mu_{XC} = a$. And the difference between the group means on the Y measure equals (a + b) where b is any constant, $\mu_{YP} - \mu_{YC} = a + b$. Then for fan spread data and equally distant time points the difference between the group means on the Z measure would equal a - b; $\mu_{ZP} - \mu_{ZC} = a - b$. The effect estimated by the adjusted gain score strategy can be written as the following:

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - 2(\mu_{xp} - \mu_{xc}) + (\mu_{zp} - \mu_{zc})$$

a + b - 2(a) + (a - b)
= 2a - 2a
= 0.

The adjusted gain score procedure does not require equal time periods between the administration of the tests. The group's growth rate can be adjusted for differences in time periods by the ratio of the time under investigation to the time between the first and second pretests. Since the adjusted gain scores are only a function of means and thus unaffected by errors of measurement, the procedure is appropriate for both the manifest and latent fan spread models. Finally the adjusted gain score strategy is not influenced by the model of within-group growth, it is therefore appropriate for both condition 1, $\rho = 1$ and condition 2, $\rho \neq 1$ of the fan spread hypothesis.

The estimate of the group difference provided by the adjusted gain score approach using the data for Table 1 is as follows:

$$\alpha_{AGS} = (\mu_{yp} - \mu_{yc}) - (\mu_{xp} - \mu_{xc}) - [(\mu_{xp} - \mu_{zp}) - (\mu_{xc} - \mu_{zc})]$$

$$(7 - 14) - (5 - 10) - [(5 - 3) - (10 - 6)]$$

$$0 = -7 + 5 - [(2 - 4)].$$

For the data found in Table 2 the estimate of the group difference is as follows:

$$\alpha_{AGS} = (\mu_{yp} - \mu_{yc}) - (\mu_{xp} - \mu_{xc}) - [(\mu_{xp} - \mu_{zp}) - (\mu_{xc} - \mu_{zc})]$$

$$(28 - 34) - (16 - 20) - [(16 - 4) - (20 - 6)]$$

$$0 = -6 + 4 - [12 - 14].$$

Both of these examples assume that the period of time between t_0 and t_1 is equal to the period from t_1 to t_2 .

Estimation With Multiple Fallible Covariates

The final analytic strategy suggested to adjust for the problem of differential growth rates between comparison groups was the analysis of covariance model with multiple covariates. The multiple covariates are the double pretest data collected prior to the period of intervention. In the review of the literature it was pointed out that because of measurement errors in the covariates, the analysis of covariance model has been judged as inappropriate in quasi-experimental studies. It was also pointed out, however, that Keesling and Wiley (1976) have suggested an approach which adjusts for the fallible covariates. As a result their technique was suggested as a possible solution to the fan spread problem. Following the Keesling-Wiley procedure, an estimate of the group differences is stated as the following:

$$\alpha_{MAC} = \mu_{yp} - \mu_{yc} - \frac{\rho_{T_x T_y} - \rho_{T_x T_z} \rho_{T_z T_z} \sigma_{T_y}}{1 - \rho_{T_x T_z}^2 \sigma_{T_x} \sigma_{T_x}} \frac{\sigma_{T_y}}{\sigma_{T_x}} (\mu_{xp} - \mu_{xc}) - \frac{\rho_{T_x T_y} - \rho_{T_x T_z} \rho_{T_x T_z} \sigma_{T_y}}{1 - \rho_{T_x T_z}^2 \sigma_{T_x} \sigma_{T_z}} \frac{\sigma_{T_y}}{\sigma_{T_z}} (\mu_{zp} - \mu_{zc}).$$

On the surface this estimate differs considerably from the one suggested by the definition of the fan spread model. The nature of the coefficients must therefore be examined. For condition 1 the relationship between the test performance was said to be perfect, $\rho = 1$. Under this assumption, then, the following is true: $\rho_{T_xT_y} = \rho_{T_xT_z} = \rho_{T_zT_y} = 1$. In the above adjustment coefficients, there appears the following factor in the denominator: $1 - \rho_{T_xT_z}^2$. If, as condition 1 suggests, the true relationship between test performance is perfect,

then denominators in these coefficients are zero, and the coefficients are undefined. The Keesling-Wiley approach is inappropriate for condition 1 when the covariates are repeated administrations of the same or parallel forms of the posttest measure.

Condition 2, on the other hand, suggests that the relationship between test scores across time is not perfect $\rho \neq 1$. Under this condition the adjustment coefficients suggested by the analysis of covariance strategy are defined. It is necessary then to determine whether or not the coefficients provide the appropriate adjustment. For the Keesling-Wiley procedure to be appropriate under condition 2 of the fan spread model, the following equality must be true:

$$\frac{\rho_{yx} - \rho_{xz}\rho_{zy}}{1 - \rho_{xz}^2} \frac{\sigma_y}{\sigma_x} (\mu_{xp} - \mu_{xc}) + \frac{\rho_{zy} - \rho_{xz}\rho_{xy}}{1 - \rho_{xz}^2} \frac{\sigma_y}{\sigma_z} (\mu_{zp} - \mu_{zc}) = \frac{\sigma_y}{\sigma_z} (\mu_{xp} - \mu_{xc})$$

Where in the above expression the correlation coefficients, (ρ) , and the variances, (σ) , are in terms of the latent true variables. The fan spread model defines the differences between the group means on the Z variable as the following:

$$\mu_{zp} - \mu_{zc} = \frac{\sigma_z}{\sigma_x} (\mu_{xp} - \mu_{xc})$$

The Keesling-Wiley coefficient can then be written as

$$\frac{\rho_{xy} - \rho_{xz}\rho_{zy}}{1 - \rho_{xz}^{2}} \frac{\sigma_{y}}{\sigma_{x}} (\mu_{xp} - \mu_{xc}) + \frac{\rho_{zy} - \rho_{xz}\rho_{xy}}{1 - \rho_{xz}^{2}} \frac{\sigma_{y}}{\sigma_{x}} (\mu_{xp} - \mu_{xc})$$

$$= \frac{\sigma_{y}}{\sigma_{x}} (\mu_{xp} - \mu_{xc}) \left[\frac{\rho_{xy} - \rho_{xz}\rho_{zy}}{1 - \rho_{xz}^{2}} + \frac{\rho_{zy} - \rho_{xz}\rho_{xy}}{1 - \rho_{xz}^{2}} \right] .$$

For the Keesling-Wiley procedure to be appropriate, then, the factor in brackets must equal unity:

$$\frac{\rho_{xy} - \rho_{xz}\rho_{zy}}{1 - \rho_{xz}^2} + \frac{\rho_{zy} - \rho_{xz}\rho_{xy}}{1 - \rho_{xz}^2} = 1.$$

This may be rewritten as the following:

$$\rho_{xy} + \rho_{zy} - \rho_{xz} (\rho_{zy} + \rho_{xy}) = 1 - \rho_{xz}^2$$

or

$$(\rho_{xy} + \rho_{zy}) (1 - \rho_{xz}) = 1 - \rho_{xz}^{2}$$

Further simplification of the equation is provided by noting that:

$$1 - \rho_{XZ}^2 = (1 - \rho_{XZ}) (1 + \rho_{XZ}).$$

Thus for the Keesling-Wiley procedure to appropriately adjust for the fan spread hypothesis under the second model of within group growth

$$(\rho_{xy} + \rho_{zy}) = (1 + \rho_{xz}).$$

This can only be true if both ρ_{xy} and ρ_{zy} are greater than ρ_{xz} . Since ρ_{zy} involves variables measured at two points farther apart in time than ρ_{xz} , the above is highly unlikely. Given the more reasonable assumption that the correlations between any two equally distant points in time are equal, the statement cannot be true. Thus the Keesling-Wiley procedure is inappropriate as a solution to the fan spread model. The data presented in Table 2 demonstrate the inappropriateness of the Keesling-Wiley procedure for condition 2 of the fan spread model. For these data the true relationships between the measures are as follows:

$$\rho_{zx}$$
 = .85, ρ_{zy} = .70, and ρ_{xy} = .90.

With these values, the estimate of the group difference is as in the following:

$$\alpha_{MAC} = (\mu_{yp} - \mu_{yc}) - \frac{\rho_{xy} - \rho_{xz}\rho_{zy}}{1 - \rho_{xz}^2} \frac{\sigma_y}{\sigma_x} (\mu_{xp} - \mu_{xc}) - \frac{\rho_{zy} - \rho_{xz}\rho_{xy}}{1 - \rho_{xz}^2} \frac{\sigma_y}{\sigma_z} (\mu_{zp} - \mu_{zc})$$

$$(28 - 34) - \frac{.90 - (.85)(.70)}{1 - (.85)^2} \frac{4.74}{3.16} (16 - 20)$$

$$- \frac{.70 - (.85)(.90)}{1 - (.85)^2} \frac{4.74}{1.56} (4 - 6)$$

$$-6 - (1.099)(1.5)(-4) - (-.2342)(3.038)(-2)$$

$$-6 + 6.595 - 1.4229$$

$$= -.8279.$$

The analysis of covariance with the two pretests as covariates does not therefore estimate the desired effect under the fan spread model for either condition 1, $\rho = 1$, or condition 2, $\rho \neq 1$.

The examination of the estimates of group differences by the four analytic strategies has shown that for condition 1, $\rho = 1$, gains in standard scores, analysis of covariance with estimated true scores of the covariate, and adjusted gain scores all estimate the effect of interest. Only the Keesling-Wiley analysis with the double pretests as covariates was shown to estimate the wrong effect. For condition 2, $\rho \neq 1$, only gains in standard scores and adjusted gain scores were shown to estimate the desired effect. Thus, researchers have a choice in selecting an analytic strategy in both condition 1 and 2 of the fan spread model. Given these findings, the selection of one technique over another might be based on precision.

Precision

The precision of an analytic strategy is defined in terms of the standard error of a simple contrast which in turn is determined by the variability of the adjusted variable. Therefore, a comparison of the standard errors associated with each analytic strategy provides a way by which the precision of each technique can be assessed.

It was pointed out in Chapter 2 that each analysis strategy being considered could be conceptualized in terms of an adjusted dependent measure. In comparing a program group with a control group, the contrast of interest is the difference between the means of the two groups on the adjusted variable. The standard error is therefore the square root of the variance of this contrast, $\sqrt{Var(W_p - W_c)}$. The variance of the contrast is defined as

$$Var (\overline{W}_{p} - \overline{W}_{c}) = Var (\overline{W}_{p}) + Var (\overline{W}_{c}) - 2 Cov (\overline{W}_{p}, \overline{W}_{c}).$$
(8)

To determine the standard error of the contrast, both the variance and the covariance of the adjusted means are needed. All of the analytic strategies considered take this general form and differ only in the approach used to define the adjusted variable, W.

Standard Error For an Index of Response

Since the gains in standard scores and the analysis of covariance model with estimated true scores as the covariate use only a single pretest and the posttest, the two approaches are similar in how the adjusted variable is created. They both take the general form of an index of response,

$$W = Y - KX$$

Where X and Y are the pretest and posttest scores, respectively, and K is the adjustment coefficient. The contrast of interest, however, is stated in terms of the group means on the adjusted variable which can be written as: $\overline{W} = \overline{Y} - K\overline{X}$. The variance of the adjusted mean is the variance of the linear combination of posttest means minus the product of the adjustment coefficient and the pretest mean,

$$Var (\overline{W}) = Var (\overline{Y} - K\overline{X})$$

$$= Var (\overline{Y}) + Var (K\overline{X}) - 2 Cov (\overline{Y}, K\overline{X}).$$
(9)

If K is known to the researcher independently of his data--that is the population parameter K is a known constant--the variance of the mean of the adjusted variable could be further simplified to:

$$Var (\overline{W}) = Var (\overline{Y}) + K^2 Var (\overline{X}) - 2K Cov (\overline{Y}, \overline{X}).$$

The strategies suggested by Kenny (gains in standard scores) and Porter (analysis of covariance with estimated true score of the covariate) do not assume knowledge of the correct adjustment coefficient; rather, they require that the coefficient be estimated on the data. As a result of fluctuation from sample to sample, the adjustment coefficient must be considered a random variable. In determining the variance of the adjusted mean, the adjustment coefficient appears only in combination with the pretest mean. Thus, to determine the variability of the adjusted variable, the variance of the product of two random variables and the covariance of a random variable and the product of two random variables are needed. Keesling and Wiley (1976) have derived statements for both of these terms assuming that K is independent of both X and Y. The variance of the product of two independent random variables can be written:

$$Var (K\overline{X}) = E(K)^2 Var (\overline{X}) + E(X)^2 Var (K) + Var (K) Var (X).$$
(10)

The covariance of a random variable with the product of two independent random variables can be written:

$$Cov (\overline{Y}, K\overline{X}) = E(K) Cov (\overline{Y}, \overline{X}),$$

where again, K is assumed to be independent of the pretest and posttest means. The proof of these two statements is provided in Appendices A and B. In summary, then, the variance of the adjusted variable is:

$$Var (\overline{W}) = Var (\overline{Y}) + E(K)^{2} Var (\overline{X}) + E(\overline{X})^{2} Var (K) + Var (K) Var (\overline{X})$$
$$- 2E(K) Cov (\overline{X}, \overline{Y}).$$

Finally, to determine the variance of the contrast of interest (equation 9), the covariance of the two adjusted means is needed; i.e., Cov $(\overline{W}_{p}, \overline{W}_{c})$. Since only the adjusted coefficient is common to the group mean on the adjusted variable, Keesling and Wiley state this covariance term as:

$$Cov (\overline{W}_{p}, \overline{W}_{c}) = E(\overline{X}_{p}) E(\overline{X}_{c}) Var (K).$$
(11)

The proof of this statement is provided in Appendix C.

Substituting equations 10 and 11 in equation 9, and assuming a balanced design with the program and control groups having equal variance on the pretest and posttest measures, the variance of the contrast between two group means on an adjusted variable is:

$$Var (\overline{W}_{p} - \overline{W}_{c}) = 2 [Var (\overline{Y}) + E(K)^{2} Var (\overline{X}) + Var (K) Var (\overline{X})$$
$$- 2E(K) Cov (\overline{X}, \overline{V})] + E(\overline{X}_{p})^{2} Var K + E(\overline{X}_{c})^{2} Var (K)$$
$$- 2E(\overline{X}_{p}) E(\overline{X}_{c}) Var (K).$$

This can be simplified to:

$$Var (\overline{W}_{p} - \overline{W}_{c}) = 2 [Var(\overline{Y}) + (E(K)^{2} + Var(K)) Var (\overline{X})$$
$$- 2E(K) Cov (\overline{X}, \overline{Y})] + [E(\overline{X}_{p}) - E(\overline{X}_{c})]^{2} Var (K).$$

Furthermore, $E(\overline{X}_p) = \mu_{xp}$ and $E(\overline{X}_c) = \mu_{xc}$, so that,

$$Var \left(\overline{W}_{p} - \overline{W}_{C}\right) = 2 \left[Var \left(\overline{Y}\right) + (E(K)^{2} + Var(K)) Var(\overline{X}) - 2E(K) Cov(\overline{Y},\overline{X})\right] + (\mu_{Xp} - \mu_{XC})^{2} Var(K).$$
(12)

Equation 12 is a general statement of the variance of the contrast of interest.

The Standard Error For Kenny's Procedure

The strategy suggested by Kenny defined the adjustment coefficient as the rato of the pooled standard deviations of the posttest to the pretest: $K = \frac{S_y}{S_x}$. To determine the expected value and the variance of this coefficient, the density function of the ratio of two correlated standard deviations is needed. Bose (1935) and Finney (1938) have derived this function from the bivariate normal density function. The density function of two correlated standard deviations is written as the following:

$$dF_{v} = \frac{2(1-\rho^{2})^{\frac{n-1}{2}}}{\beta(\frac{n-1}{2},\frac{n-1}{2})} \frac{v^{n-2}}{(1+v^{2})^{n-1}} \left\{ 1 - \frac{4\rho^{2}v^{2}}{(1+v^{2})^{2}} \right\}^{-\frac{n}{2}}$$

where:

- represents the relationship between the pretest and posttest for the population on the manifest variables;
- β is the beta function;

$$v = \frac{S_y}{\sigma_y} / \frac{S_x}{\sigma_x}$$
; and

n is the sample size.

The expected value of the ratio of the standard deviation is determined over the interval from 0 to ∞ the density function as in the following:

$$E(v) = \int_0^\infty v \, dF_v.$$

To determine the variance of the coefficient, the expected value of the squared ratio of the standard deviation is needed. This value is calculated by integrating from 0 to ∞ the product of the ratio of the variances and the density function:

$$E(v^2) = \int_0^\infty v^2 dF_v$$

Solving the above equation results in:

$$E\left(\frac{S_y^2}{S_x^2}\right) = \frac{(n-1)-2\rho^2}{n-3} \frac{\sigma_y^2}{\sigma_x^2} .$$

Appendix D provides the calculation leading to the above solution. The solution to determining the expected value of the ratio of standard deviation was more difficult and required the numerical integration of the following function:

$$E\left(\frac{S_{y}}{S_{x}}\right) = \frac{\Gamma(n-1)}{\left[\Gamma\left(\frac{n-1}{2}\right)\right]^{2}} \int_{0}^{1.5} \sqrt{(1-\rho^{2}) + 4\rho^{2} Y(1-Y)} \left(Y(1-X)\right)^{\frac{n-4}{2}} dy \frac{\sigma_{y}}{\sigma_{x}}.$$

The derivation of this solution is provided in Appendix E. To integrate this function, a "canned" computer program called DCADRE (de Boor, 1971) was used in conjunction with the CDC6500 computer at Michigan State University. The solution of the equation for values of n ranging from 20 to 120 in multiples of 10 and values of ρ for the manifest variables ranging from .1 to .9 in increments of .1 are presented in Table 3. Earlier in this chapter it was stated that the ratio of sample standard deviation was a biased estimator of the rate of population standard deviation. The results found in Table 3 confirm that statement. If the ratio of the standard deviations was an unbiased estimator of the ratio of the population standard deviations then all entries in Table 3 would have been one. For a sample of size 30 and a correlation of .1 between the measures, the ratio of the standard deviations equals 1.0178. As the sample size becomes large and the correlation between measures high, the magnitude of the bias is reduced. For example, with a sample of 50 and a correlation of .8, the ratio of the sample standard deviations equal 1.0038. Thus the procedure suggested by Kenny provides the appropriate estimate of the group differences in a fan spread situation only when the sample is large and the correlation between measures is high.

The variance of the adjustment coefficient suggested by Kenny is determined by:

$$\operatorname{Var}\left(\frac{S_{\mathbf{y}}}{S_{\mathbf{x}}}\right) = \left[E\left(\frac{S_{\mathbf{y}}^{2}}{S_{\mathbf{x}}^{2}}\right) - \left(E\left(\frac{S_{\mathbf{y}}}{S_{\mathbf{x}}}\right)\right)^{2} \right].$$

These results are presented in Table 4, assuming the population variances of the two measures are equal. For a sample of 30 and a correlation of .1, the variance of the ratio of the standard deviations equals .03733. As the sample size and the correlation

The Expected Value of the Ratio of Two Non-Independent Sample Standard Deviations When the Population Standard Deviations Are Equal

NX L	۲.	.2	.3	.4	.5	.6	.7	8.	6.
20	1.0279	1.0271	1.0257	1.0238	1.0213	1.0183	1.0146	1.0104	1.0055
30	1.0178	1.0173	1.0164	1.0152	1.0136	1.0116	1.0093	1.0066	1.0035
40	1.0131	1.0127	1.0121	1.0112	1.0100	1.0085	1.0068	1.0048	1.0026
50	1.0104	1010.1	1.0095	1.0088	1.0079	1.0067	1.0054	1.0038	1.0020
60	1.0086	1.0083	1.0079	1.0073	1.0065	1.0056	1.0044	1.0031	1.0017
70	1.0073	1.0071	1.0067	1.0062	1.0056	1.0048	1.0038	1.0027	1.0014
80	1.0064	1.0062	1.0059	1.0054	1.0048	1.0041	1.0033	1.0023	1.0012
06	1.0056	1.0055	1.0052	1.0048	1.0043	1.0037	1.0029	1.0021	1.0011
100	1.0051	1.0049	1.0047	1.0043	1.0038	1.0033	1.0026	1.0019	1.0010
011	1.0046	1.0045	1.0042	1.0039	1.0025	1.0030	1.0024	1.0017	1.0009
120	1.0042	1.0041	1.0039	1.0036	1.0032	1.0027	1.0022	1.0015	1.0008

Ð
_
Δ
Ø
F

The Variance of the Ratio of Two Non-Independent Sample Standard Deviations When the Population Standard Deviations Are Equal

n pxy		.2	е.	4.	.5	.6	۲.	ω.	6.
20	.05992	.05807	.05498	.05066	.04513	.03840	.03049	.02143	.01125
30	.03733	.03619	.03428	.03160	.02817	.02400	.01908	.01343	.00706
40	.02712	.02628	.02490	.02296	.02048	.01745	.01388	.00978	.00515
50	.02129	.02064	.01955	.01803	.01609	.01371	16010.	.00769	.00405
60	.01752	.01698	.01609	.01485	.01325	.01129	.00899	.00634	.00334
70	.01484	.01439	.01363	.01258	.01122	.00956	.00760	.00535	.00280
80	.01297	.01253	.01188	.01096	.00978	.00833	.00663	.00468	.00246
06	.01144	60110 .	.01051	00670.	.00866	.00738	.00588	.00414	.00218
100	.01026	.00995	.00948	.00870	.00761	.00662	.00527	.00372	.00196
011	.00930	10600.	.00854	.00788	.00703	.00600	.00478	.00337	.00178
120	.00850	.00824	.00781	.00721	.00643	.00549	.00437	.00308	.00163

between the measures increase, the variance of the ratio decreases. For a sample of 50 and a correlation of .8 the variance of the ratio of the standard deviations equals .00769.

The Standard Error For Porter's Procedure

The analytic strategy suggested by Porter using the analysis of covariance model with estimated true score of the covariate, defines the adjustment coefficient as the ratio of the linear regression slope $K = \frac{D_{y \cdot x}}{P_{v \cdot x}} .$ of posttest on pretest to the reliability of the covariate: One way of calculating the reliability coefficient is to use the sample data and estimate a measure of internal consistency for the pretest. It was suggested earlier, however, that a test-retest reliability was a more desirable coefficient. If one of the two administrations is also used as the covariate, there would be a lack of independence between the regression slope and the reliability coefficient, thus complicating the estimation of the expected value and the variance of the adjustment. A second approach of estimating the reliability coefficient would be to select an independent sample of subjects and administer the instrument twice. This estimate of the test-retest reliability could be obtained in the pilot testing of the instrumentation. In both of these methods the resulting estimate of the reliability coefficient is a random variable whose fluctuation from sample to sample must be taken into consideration in estimating the variance of the adjustment coefficient. A third way of determining the reliability coefficient is to use the published test-retest reliability coefficient if one is available. The researcher, however, must be careful that the reported reliability is appropriate for the population he is studying. This latter approach was taken in the present study because it seemed likely that a standardized instrument would be used in an evaluation study and because it simplified calculations since the coefficient would be constant across replications.

Given that the reliability estimate is known, the expected value of the adjustment coefficient is the product of the reciprocal of the reliability coefficient and the expected value of the regression slope:

$$E\left(\frac{b_{y*x}}{\rho_{xx}}\right) = \frac{1}{\rho_{xx}} E(b_{y*x}),$$

and the variance is:

$$\operatorname{Var}\left(\frac{b_{y*x}}{\rho_{xx}}\right) = \frac{1}{\rho_{xx}^2} \operatorname{Var}\left(b_{y*x}\right).$$

Thus the problem of determining the expected value and variance of the adjustment coefficient is simplified to determining the expected value and variance of the regression slope. To obtain these values, the density function of a regression coefficient is needed. This function is well known and appears in several forms. Kendall and Stewart (1952) present the density function of the regression coefficient as:

$$dF_{\beta} = \frac{\Gamma\left(\frac{n}{2}\right)(1-\rho^{2})\frac{n-1}{2}}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)} \left(\frac{\sigma_{y}}{\sigma_{x}}\right)^{n-1} \frac{db}{\left\{\frac{\sigma_{z}^{2}}{\sigma_{1}}(1-\rho^{2}) + \left(b-\frac{\rho\sigma_{y}}{\sigma_{x}}\right)^{2}\right\}^{n/2}}$$

Where

 ρ, σ_y, σ_x , and n are as defined previously; and Γ represents the gamma function.

The expected value of the regression slope is determined by integrating over the range, $-\infty$ to $+\infty$, the product of the regression coefficient and the density function,

$$E(b_{y*x}) = \int_{-\infty}^{+\infty} b dF_{\beta}.$$

The solving of this integral results in:

$$E(b_{y \cdot x}) = \rho_{yx} \frac{\sigma_y}{\sigma_x}.$$
 (13)

Appendix F presents the details of the integration. This result shows that the sampled regression coefficient is an unbiased estimator of the population regression coefficient. To determine the variance of the regression slope, the expected value of the squared regression coefficient is needed. This is determined by integrating over the range of $-\infty$ to $+\infty$ the product of the squared regression coefficient and the density function:

$$E(b_{y^*X}^2) = \int_{-\infty}^{+\infty} b^2 dF_{\beta}.$$

Appendix G contains the solution to this integration problem.

$$E(b_{y \cdot x}^{2}) = \frac{(1 - \rho_{yx}^{2})}{n - 3} \frac{\sigma_{y}^{2}}{\sigma_{x}^{2}} + \rho_{yx}^{2} \frac{\sigma_{y}^{2}}{\sigma_{x}^{2}}.$$
 (14)

Combining equations 13 and 14:

Var
$$(b_{y \cdot x}) = E(b_{y \cdot x}^2) - (E(b_{y \cdot x}))^2$$
,
= $\frac{(1 - \rho^2)}{n - 3} \frac{\sigma_y^2}{\sigma_x^2}$.

In the earlier discussion of the adjustment coefficient suggested by Kenny, it was assumed that the standard deviations of the two variables were equal. Making the same assumption here simplifies the formula for computing the variance of the regression coefficient and facilitates comparisons with Kenny's adjustment coefficient. The variance of the regression coefficient therefore is equal to:

Var
$$(b_{y \cdot x}) = \frac{1 - \rho^2}{n - 3}$$

Table 5 contains values of this ratio for the same values of ρ and n that were considered in the discussion of Kenny's adjustment coefficient. For example, the variance of the regression slope with a sample size of 30 and a correlation between the measures equaling .1 is equal to .03667. As the sample size and the correlation between the measures increase, the variance of the regression slope decreases. With a sample of 50 and a correlation of .8, the variance of the regression slope equals .00766. The variance of the adjustment coefficient suggested

ے ا	
ab	
F	

The Variance of a Sample Regression Coefficient When the Population Standard Deviations of the Two Variables Are Equal

y ry	۲.	.2	.3	.4	.5	.6	٦.	8.	6.
20	.05824	.05647	.05353	.04941	.04412	.03765	.03000	.02118	.01118
30	.03667	.03556	.03370	.03111	.02778	.02370	.01889	.01333	.00704
40	.02676	.02595	.02460	.02270	.02027	.01730	.01378	.00973	.00514
50	.02106	.02043	.01936	.01787	.01596	.01362	.01085	.00766	.00404
60	.01737	.01684	.01597	.01474	.01316	.01123	.00895	.00632	.00333
70	.01478	.01433	.01358	.01254	61110.	.00955	.00761	.00537	.00284
80	.01286	.01247	.01182	16010.	.00974	.00831	.00662	.00468	.00247
06	.01138	.01103	.01046	.00966	.00862	.00736	.00586	.00414	.00218
100	.01021	06600.	.00938	.00866	.00773	.00660	.00526	.00371	96100.
011	.00925	.00897	.00850	.00785	.00701	.00598	.00477	.00336	.00178
120	.00846	.00821	.00778	.00718	.00641	.00547	.00436	.00308	.00162

by Porter can then be calculated by multiplying the variance of the regression slope by the reciprocal of the reliability coefficient squared. The reliability coefficient for each entry in Table 5 equaled the respective correlation between the measures. These results are presented in Table 6. For a sample of size 30 and a correlation between the measures equaling .1 with a reliability coefficient of .1, the variance of the adjustment coefficient equals 3.6667. As the sample size increases and the correlation and reliability increase, the variance of the adjustment coefficient decreases. For a sample of 50 and a correlation of .8 with a reliability coefficient of .8, the variance of the adjustment coefficient equals .01197.

The Standard Error For Adjusted Gain Scores

The derivation of the standard error for the adjusted gain score strategy involves a slightly different approach. An individual's adjustment score is determined by the following formula:

 $W = Y - X - (\overline{X} - \overline{Z})$

or

```
W = Y - X - \overline{X} + \overline{Z}.
```

Where

Z,X, and Y are the first pretest, the second pretest (administered just prior to the introduction of the treatment) and the posttest, respecitvely.

The variance of this variable can be written as the following:

9	
Ð	
lď	
Ē	

The Variance of the Adjustment Coefficient Suggested in the True Score Analysis of Covariance Procedure

XX) u		.2	.3	.4	.5	.6	.7	8.	6.
20	5.8235	1.41178	.59477	.30883	.17647	.10458	.06122	.03309	.01379
30	3.6667	.88890	.37449	.19444	.11112	.06584	.03855	.02083	.00869
40	2.6757	.64865	.27378	.14189	.08108	.04805	.02813	.01520	.00634
50	2.1064	.51065	.21513	0/111.	.06383	.03783	.02215	76110.	.00499
60	1.7368	.42105	.17740	11260.	.05263	.03119	.01826	.00987	.00412
70	1.4776	.35820	.15091	.07836	.04478	.02653	.01553	.00840	.00350
80	1.2857	.31170	.13202	.06818	.03896	.02309	.01360	.00731	.00305
06	1.1379	.27585	.11622	.06035	.03448	.02045	.01196	.00647	.00269
100	1.0206	.24742	.10424	.05412	.03093	.01833	.01073	.00580	.00242
110	0.92523	.22430	.04450	.04907	.02804	.01662	.00973	.00525	.00220
120	0.84615	.20513	.08642	.04487	.02564	.01520	.00890	.00481	.00201

Var (W) = Var (Y) + Var (X) + Var (\overline{X}) + Var (\overline{Z}) - Cov (Y,X) - 2 Cov (Y, \overline{X})

+ 2 Cov
$$(Y,\overline{Z})$$
 + Cov (X,\overline{X}) - 2 Cov (X,\overline{Z}) - 2 Cov $(\overline{X},\overline{Z})$.

The above can be simplified to:

Var (W) = Var (Y) +
$$(1 + \frac{3}{n})$$
 Var (X) - $(2 + \frac{2}{n})$ Cov (X,Y)
+ $\frac{1}{n}$ [Var (Z) + 2 Cov (Y,Z) - 4 Cov (X,Z)]

(see Appendix H). To facilitate comparisons with the two previous standard errors, the above variance of the contrast of adjusted group means can be written as the following:

$$Var (\overline{W}) = Var (\overline{Y}) + (1 + \frac{3}{n}) Var (\overline{X}) - (2 + \frac{2}{n}) Cov (\overline{X}, \overline{Y})$$
$$+ \frac{1}{n} [Var (\overline{Z}) + 2 Cov (\overline{Y}, \overline{Z}) - 4 Cov (\overline{X}, \overline{Z})],$$
$$= \frac{1}{n} [Var (Y) + (1 + \frac{3}{n}) Var (X) - (2 + \frac{2}{n}) Cov (X, Y)]$$
$$+ \frac{1}{n^2} (Var (Z) + 2 Cov (Y, Z) - 4 Cov (X, Z) .$$

The Standard Error For the Keesling-Wiley Procedure

Since the analysis of covariance with multiple covariates was shown to be inappropriate in fan spread situations for either condition 1 or 2, the standard error associated with that technique is not considered.

To present: analyti the sta

determi

compone

of the

include

of cova

between

compone

the fir

second

are det

precisi

the coe

Tł

^{three} a ^{Varian(}

differ

^{score} a ^{Using} 1

^{COeffic}

^{vided} i

^{other} h

To facilitate comparisons of statistical precision, Table 7 presents in summary, the standard errors associated with the competing analytic strategies. The formulas presented in Table 7 indicate that the standard errors of the analytic strategies being considered are determined by a combination of four components. The first three components involving: the variance of the posttest, the variance of the pretest, and the covariance of the pretest and posttest, are included in all three standard errors. Standardized gains and analysis of covariance with estimated true scores consider the squared difference between the population means in the fourth component, while the fourth component of the adjusted gain score strategy involves the variance of the first pretest and the covariance of the first pretest with the second pretest and with the posttest. Since the three standard errors are determined by basically the same components, the differences in the precision of the analytic strategies can be explained by differences in the coefficients of the four components.

The first component of each standard error is identical for all three analytic strategies with 1 as the coefficient of the posttest variance term. The coefficient of the last three components, however, differ considerably. Both gains in standard scores and estimated true score analysis of covariance strategies determine these coefficients using the expected value and variance of their respective adjustment coefficients. These expected values and variance estimates were provided in Tables 3, 4, and 6. The adjusted gain score approach, on the other hand, determines the coefficient of the last three components

 $\frac{2}{n}\left[\operatorname{Var}\left(Y\right) + \left[\frac{\left(\mathsf{E}(\mathsf{b}_{\mathbf{y},\mathbf{x}})\right)^2}{\rho_{\mathbf{x}\mathbf{x}}^2} + \frac{\operatorname{Var}\left(\mathsf{b}_{\mathbf{y},\mathbf{x}}\right)}{\rho_{\mathbf{x}\mathbf{x}}^2}\right]\operatorname{Var}\left(X\right) - 2\frac{\mathsf{E}(\mathsf{b}_{\mathbf{y}},\mathbf{x})}{\rho_{\mathbf{x}\mathbf{x}}}\operatorname{Cov}\left(X,Y\right)\right] + \left(\mu_{\mathbf{x}p} - \mu_{\mathbf{x}c}\right)^2 \frac{\operatorname{Var}\left(\mathsf{b}_{\mathbf{y},\mathbf{x}}\right)}{\rho_{\mathbf{x}\mathbf{x}}^2}$ $\left| \frac{2}{n} \left| \text{Var}\left(Y \right) + \left[\left(E \left(\frac{S}{\delta_x} \right) \right)^2 + \text{Var}\left(\frac{S}{\delta_x} \right) \right] \text{Var}\left(X \right) - 2E \left(\frac{S}{\delta_x} \right) \text{Cov}\left(X, Y \right) \right] + \left(\mu_{Xp} - \mu_{Xc} \right)^2 \quad \text{Var}\left(\frac{S_Y}{\delta_x} \right) \right) + \left(\frac{S}{\delta_x} \right) \left(\frac{S$ $\left| \frac{2}{n} \left| \text{Var}(Y) + \left(1 + \frac{3}{n} \right) \text{Var}(X) - \left(2 + \frac{2}{n} \right) \text{Cov}(X, Y) \right| + \frac{2}{n^2} \left| \text{Var}(Z) + 2 \text{ Cov}(Y, Z) - 4 \text{ Cov}(X, Z) \right| \right|$ Standard error for analysis of covariance for estimated true scores of the covariate: The Standard Errors Associated With Gains in Standard Scores, True Score Analysis of Covariance, and Adjusted Gain Scores Standard error for gains in standard scores: Standard error for adjusted gain scores:

Table 7

solely on the sample size. Tables 8 through 13 present the coefficients associated with the second, third, and fourth components of the standard errors for the three analytic strategies being considered. For example, the standard error for each competing analytic strategy when the sample size equals 60 and the correlation between the measures equals .8 (see Table 9) can be computed as the following:

for gains in standard errors:

$$\sqrt{\frac{2}{60}} \left[Var(Y) + (1.0194) Var(X) - 2.0096 Cov(X,Y) + .00634 (\mu_{Xp} - \mu_{Xc})^2 \right]$$

for true score analysis of covariance:

$$\sqrt{\frac{2}{60}} \left[Var(Y) + (1.00968) Var(X) - 2.0 Cov(X,Y) \right] + .00968 (\mu_{xp} - \mu_{xc})^2$$

for adjusted gain scores:

Tables 8 through 13 were compiled using the previously derived expected values and variances of the adjustment coefficients when the variance of the pretest and posttest are equal. The fan spread model, however, suggests increasing variability from pretest to posttest. This assumption has the effect of increasing the expected value and variance of the adjustment coefficients (see Tables 3, 4, and 6) by a factor of the population posttest to pretest variance ratio, $\frac{\sigma_y^2}{\sigma_y^2}$. In general,

	Coefficients for th Strategies Wh	e Second, Third, and Four en ρ = .9 for the Manifes	th Components t Variables a	of the Star Ind the Popul	ldard Error As lation Varianc	ssociated With ce of the X ar	i the Three Id Y Variabl	Competing Anal es Are Equal	lytic
	Second	component Var (X)		Third	component Cov	(X,Y)	Fc	ourth component	
	Gains in standard scores	Analysis of covariance	Adjusted gain scores	Gains in standard scores	Analysis of covariance	Adjusted gain scores	Gains in standard scores	Analysis of covariance	Adjusted gain scores
c	$\left[\mathbb{E} \left(\frac{S_{y}}{S_{x}} \right) \right]^{2} + \sigma_{S_{y}/S_{x}}^{2}$	$\frac{1}{\rho_{xx}^2} \left[E(b_{y,x})^2 + \sigma_{b_{y,x}}^2 \right]$	∞ ⊑ + Γ	$2E\left(\frac{5}{5x}\right)$	2 ^{E(b} y.x) ^p xx	2 + <u>2</u>	σ_{S_y/S_X}^2	σ ² _{by•x} /ρ ² xx	2 12
20	1.02228	1.01380	1.15000	2.01100	2.00000	2.10000	.01125	.01380	.00500
40	1.01036	1.00634	1.07500	2.00520	2.00000	2.05000	.00514	.00634	.00125
60	1.00674	1.00411	1.05000	2.00340	2.00000	2.03333	.00334	.00411	.00055
80	1.00486	1.00428	1.03800	2.00240	2.00000	2.02500	.00246	.00428	.00031
100	1.00396	1.00242	1.03000	2.00200	2.00000	2.02000	.00196	.00242	.00020

	זרו מרכא ובא שוב		L VALIANICS AI	וח מוב נסחחוס	ILIUI VALIAILA			באתם בלחמו	
	Second	component Var (X)		Third	component Cov	(X,Y)	• 6	wrth componen	
E	$\left[E\left(\frac{S}{S_{X}}\right) \right]^{2} + \sigma_{S_{Y}/S_{X}}^{2}$	Analysis of covariance $\frac{1}{\rho_{XX}^2} \left[E(b_{y,x})^2 + \sigma_{b_{y,X}}^2 \right]$	Adjusted gain scores 1 + $\frac{3}{n}$	$\begin{array}{c} \text{Gains in}\\ \text{standard}\\ \text{scores}\\ 2E\left(\frac{S_{\chi}}{S_{\chi}}\right) \end{array}$	Analysis of covariance E (b _{y·x}) 2 P _{xx}	Adjusted gain scores 2 + $\frac{2}{n}$	Gains in standard scores ° _{Sy} /S _x	Analysis of covariance d ² ,/ ^p xx	Adjusted gain scores n ²
20	1.04233	1.03309	1.15000	2.02080	2.00000	2.10000	.02143	.03309	.00500
4	1.01940	1.01520	1.07500	2.00960	2.00000	2.05000	.00978	.01520	.00125
60	1.01255	1.00968	1.05000	2.00620	2.00000	2.03333	.00634	.00968	.00055
80	1.00428	1.00731	1.03800	2.00460	2.00000	2.02500	.00468	.00731	.00031
1 00	1.00752	1.00580	1.03000	2.00380	2.00000	2.02000	.00372	.00580	.00020

Coefficients for the Second. Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Stratecies When o = .8 for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal

Table 9

	Second	component Var (X)		Third	component (X,	۲)	Ĕ	urth componer	
	Gains in standard scores [/S./] ²	Analysis of covariance	Adjusted gain scores	Gains in standard scores / S. /	Analysis of covariance E (b)	Adjusted gain scores	Gains in standard scores	Analysis of covariance	Adjusted gain scores
E	$\left[E\left(\frac{x}{S_{x}}\right) \right]^{+\sigma_{S_{y}}^{2}/S_{x}}$	$\frac{1}{p_{XX}^2} \left[E(b_{y,\cdot,X})^2 + \sigma_{b_{y,\cdot,X}}^2 \right]$	0 = + -	$2E\left(\frac{5^{4}}{3}\right)$	2 y x x	5 + <mark>1</mark>	σ ² y/s _x	σ ² by.x ^{/p} xx	<mark>13</mark>
20	1.05489	1.06122	1.15000	2.02420	2.0000	2.10000	.03049	.06122	.00500
40	1.02753	1.02813	1.07500	2.01360	2.00000	2.05000	.01388	.02813	.00125
60	1.01781	1.01826	1.05000	2.00880	2.00000	2.03333	.00898	.01826	.00056
80	1.01324	1.01352	1.03800	2.00660	2.00000	2.02500	.00663	.01352	.00031
100	1.01048	1.01073	1.03000	2.00520	2.00000	2.02000	.00527	.01073	.00020

Coefficients for the Second. Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When o = .7 for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal

Table 10

	orrated ies when		L VALIADIES A	in rue robnia	LUNI YATIANU	י טו רוויב א מוו	I VALIADIA	s Are cquai	
	Second c	component Var (X)		Third	component Cov	(X,Y)		ourth compone	nt
	Gains in standard scores	Analysis of covariance	Adjus ted gain scores	Gains in standard scores	Analysis of covariance	Adjusted gain scores	Gains in standard scores	Analysis of covariance	Adjusted gain scores
c	$\left[E\left(\frac{3y}{5x}\right) \right] + \sigma_{S_{y}/S_{x}}^{2}$	$\frac{1}{\rho_{XX}^2} \left[E(b_{Y X})^2 \right] + \sigma_{b_{Y \cdot X}}^2$	1 + <mark>3</mark>	2 $E\left(\frac{3}{5}\chi\right)$	2 ^{E (b} y.x) ^p xx	2 + <u>2</u>	^o S _y /S _x	σ ² by.x ² xx	2 12
20	1.07533	1.10458	1.15000	2.03660	2.00000	2.10000	.03839	.10458	.00500
40	1.03452	1.04805	1.07500	2.01700	2.00000	2.05000	.01745	.04805	.00125
60	1.02252	1.03119	1.05000	2.01120	2.00000	2.03333	.01129	.03119	.00056
8	1.01655	1.02309	1.03800	2.00820	2.00000	2.02500	.00833	.02309	.00031
1 00	1.01323	1.01833	1.03000	2.00660	2.0000	2.02000	.00662	.01833	.00020

nalytic	-
A gu	Equa
mpeti	Are
e Co	bles
Thre	Varia
the	ہے P
With	Xan
ated	the
soci	e of
r As	fanc
Erro	ı Var
dard	ation
Stan	Indo
the	the P
s of	and
nent	les
Compo	ariab
rth	st Vi
1 Fou	nife
, and	э Т
htm	or t
nd, 1	9
Seco	d
the	When
for	gies
ficients	Strate
Set	

Coefficients for the Second. Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When ρ = .5 for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal

	Second c	omponent Yar (X)		Third	component Cov	(X,Y)	Fc	ourth componen	4
c	$\begin{bmatrix} Gains in \\ standard \\ scores \\ E\left(\frac{S_{y}}{S_{y}}\right)^{2} + \sigma_{S_{x}}^{2}$	Analysis of covariance $\frac{1}{\sigma^2} \left[E(b_{y_x})^2 + \sigma_{b_{xyy}}^2 \right]$	Adjusted gain scores 1 + $\frac{3}{n}$	Gains in standard scores $2E\left(\frac{S}{5^{V}}\right)$	Analysis of covariance 2 (by.x)	Adjusted gain scores 2 + $\frac{2}{n}$	Gains in standard scores $\sigma^2_{S_*}/S_*$	Analysis of covariance db/p ² xx	Adjusted gain scores n ² n ²
20	x Y L'X 'J 'I.08818	ل×۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۷ - ۲۰۰۸	1.15000	2.04260	2.00000	2.10000	y x .04513	y•x .17647	.00500
4 0	1.04058	1.08108	1.07500	2.02000	2.0000	2.05000	.02048	.08108	.00125
60	1.02628	1.05432	1.05000	2.01300	2.00000	2.03333	.01325	.05432	.00056
80	1.01939	1.03896	1.03800	2.00960	2.00000	2.02500	.00978	.03896	.00031
100	1.01523	1.03093	1.03000	2.00760	2.0000	2.02000	.007613	.03093	.00020

Coefficients for the Second, Third, and Fourth Components of the Standard Error Associated With the Three Competing Analytic Strategies When p = .4 for the Manifest Variables and the Population Variance of the X and Y Variables Are Equal

	Second	component Var (X)		Third	component Cov	(X,Y)	Fo	urth componen	t
	Gains in standard scores	Analysis of covariance	Adjusted gain scores	Gains in standard scores	Analysis of covariance	Adjusted gain scores	Gains in standard scores	Analysis of covariance	Adjusted gain scores
E	$\left[E\left(\frac{S_{y}}{S_{x}}\right) \right]^{2} + \sigma_{S_{y}/S_{x}}^{2}$	$: \frac{1}{\rho_{xx}^2} \left[E(b_{y,x})^2 + \sigma_{b_y x}^2 \right]$	1 + <mark>1</mark>	$2E\left(\frac{S}{S_{X}}\right)$	2 ^{E (b} y.x) ^p xx	$-2+\frac{2}{n}$	σ ² _S /S _X	$\sigma^2_{b_{y,x}}/\rho^2_{xx}$	2 n ²
20	1.09883	1.30883	1.15000	2.04760	2.00000	2.10000	.05066	.30883	.00500
40	1.04549	1.14189	1.07500	2.02240	2.00000	2.05000	.02296	.11419	.00125
60	1.02949	1.09211	1.05000	2.01460	2.00000	2.03333	.01485	11260.	.00056
80	1.02179	1.06818	1.03800	2.01080	2.00000	2.02500	.01096	.06818	.00031
1 00	1.01732	1.05412	1.03000	2.00890	2.00000	2.02000	.00869	.05412	.00020
									,
th st is i m S 0 ١ the greater the difference between the two variances, the larger the standard error. This results from the fact that the second component is increased by the ratio of the variances while the third component is increased by the ratio of the standard deviations. However, the magnitude of the increased standard error is the same for the gains in standard score approach as that for the estimated true score analysis of covariance technique. Thus, in comparing the precision associated with these two strategies, the coefficients found in Tables 8 through 13 provide a reasonable basis on which judgments can be made. The adjusted gain score approach is not affected by the fan spread assumption. Thus, the standard errors presented in Tables 8 through 13 remain the same regardless of the difference between the variance of the pretest and posttest. The implications of this result when comparing the standard error associated with the adjusted gain score approach with either gains in standard scores or the estimated true score analysis of covariance techniques are discussed in the following chapter.

CHAPTER 5

DISCUSSION

The previous chapter has considered the fan spread model from two perspectives. One approach, the traditional perspective, is viewed within group growth as linear with a common starting point and a constant rate of growth for an individual but varying rates of growth among individuals. This view of growth suggested a perfect relationship, $\rho = 1$, between test performances except for measurement errors. The second perspective of the fan spread model viewed within group growth as non-linear for an individual, but linear for the average of the group. This approach suggested a less than perfect relationship, $\rho \neq 1$, between test performances even with perfectly reliable measures.

In considering the traditional conceptualization of the fan spread model, gains in standard scores, estimated true score analysis of covariance, and adjusted gain score strategies, were all shown to be appropriate techniques. Choosing one of these strategies in preference to the others must therefore be based on some criterion other than the effect estimated. A second criterion was suggested involving the precision provided by each procedure in testing the common hypothesis. To evaluate the precision, the standard error associated with each approach was derived. Examining these standard errors showed that they were determined by the same basic components. The three competing strategies

have the same first three components consisting of the posttest variance, the pretest variance and the pretest-posttest covariance. They differ, however, in the coefficients associated with the second and third components. The coefficients for the gains in standard scores and true score analysis of covariance are determined by the expected value and variance of their respective adjustment coefficient. The second and third coefficients for the adjusted gain score procedure. on the other hand, are determined solely on the basis of sample size. There are considerable differences among the three strategies in the fourth component of the standard errors. The fourth component for the gains in standard scores and true score analysis of covariance is determined by the square of the population mean difference between the comparison groups on the pretest measure. They differ only in the coefficient for this component which is determined by the variance of their respective adjustment coefficients. The fourth component for the adjusted gain score procedure, on the other hand, is determined by the variance of the first pretest, the covariance of the first pretest and second pretest, and the covariance of the first pretest and the posttest. The coefficient for this component is determined by the sample size. Although there were differences in the fourth component, it was suggested that a comparison of the component coefficients could determine differences in the precision provided by the three analytic strategies. The coefficients of those components were presented in Tables 8 through 13. These coefficients were appropriate for the special case where the variance of the pretest measure and the posttest

measure were equal. Conclusions based on this data, however, can be extended to situations where the variances are unequal.

The first three tables, Tables 8, 9, and 10, present the coefficients for the situations when the correlations between measures is high ($\rho \ge .7$) which is likely for condition 1. Under this condition, there appear to be only minor differences between the coefficients defined by the three strategies for the three components. Furthermore, this result is consistent for both small and large samples. Special attention might be given to the coefficients for the fourth component. When the relationship between the measures is high, the coefficients associated with the fourth component appear to be very small. For practical purposes these coefficients could be judged to be essentially zero.

The second three tables, Tables 11, 12, and 13, present the coefficients for the situations when the correlation between measures is low, ($\rho < .7$). Under this condition, typical for condition 2 of the fan spread model, the coefficients associated with the second and third components again appear similar for the three strategies. As the sample size increases, the magnitude of the coefficients decreases, thus reducing the standard error and increasing the precision of the test. The coefficients associated with the fourth component, however, can no longer be judged as being equal across the three competing strategies. The coefficients for the fourth component of the adjusted gain score strategy are unaffected by the relationship between the measures and thus, remain essentially zero. The coefficients for gains in standard

scores and estimated true score analysis of covariance, on the other hand, are inversely related to the relationship between the measures. The effect of this relationship is greatest in small samples. Thus, in comparing the precision associated with the three strategies, the adjusted gain score procedure must be judged as providing the smallest standard error when the relationship between the measures is low. It should also be remembered that the adjusted gain score procedure is the only one of the three to provide unbiased estimates of effects for data with low correlations.

The above discussion was based on the coefficients presented in Tables 8 through 13. These values were determined for the situation when the variance of the pretest equals the variance of the posttest. The fan spread model suggests, however, that the variance increases with time. This assumption of the fan spread model has no effect on the standard error associated with the adjusted gain score strategy. The standard errors of the gains in standard scores and true score analysis of covariance are, however, affected by increasing variance. As variance increases, the expected value and variance of the adjustment coefficients increase, which in turn increases the coefficients of the three components previously discussed. In comparing the three competing analytic strategies, under increasing variance greater precision is achieved through the adjusted gain score procedure than either gains in standard scores or true score analysis of covariance. Furthermore, the difference in precision increases as: the relationship between the pretest and posttest measures decreases, the sample size decreases, and the greater fan spread.

The adjusted gain score procedure requires the availability of two pretests prior to the intervention. When these data are not obtainable, the decision must be made between gains in standard scores and estimated true score analysis of covariance. Concentrating on the coefficients for the gains in standard scores and estimated true score analysis of covariance, a review of Tables 8 through 13 indicates almost no differences when the relationship is high between the pretest and posttest measures. Sample size has very little effect on the magnitude of these coefficients in the first three tables. As the relationship decreases between the measures, the coefficients associated with the estimated true score analysis of covariance procedure became slightly larger than those for the gains in standard scores approach. This difference was more prominent in small rather than large samples. The dissimilarity between the coefficients was especially salient in comparisons considering the fourth component. These results indicate that in situations where the relationship is low between the pretest and posttest measures, gains in standard scores are likely to have a smaller standard error than estimated true score analysis of covariance. The observations on which this conclusion was based, were derived for the special case when the variance of the pretest and posttest measures was equal. In situations conforming to the fan spread model, the increase in variability affects both strategies equally. Thus, the above conclusion applies equally to situations conforming to the general fan spread hypothesis.

It should be noted that the standard error given for the gains in standard scores procedure differs from that suggested by Kenny. Kenny has suggested a two-stage process: first, determine the adjusted variable; and second, use it as the dependent variable in the analysis of variance model. This procedure assumes that the adjustment coefficient determined in step 1 is a constant across replications. Based on this assumption the standard error takes the following form:

$$\sqrt{\frac{2}{n} \left[\text{Var (Y)} + \frac{S_y^2}{S_x^2} \text{ Var (X)} - 2 \frac{S_y}{S_x} \text{ Cov (X,Y)} \right]}$$

Kenny's standard error, therefore, differs from the correct standard error presented in Table 7 by eliminating all factors involving the variability of the adjustment coefficient. This reduced form of the standard error produces spurious precision which leads to a liberal test of the hypothesis under investigation.

The degree to which Kenny's procedure is too liberal is dependent on the variability of the adjustment coefficient. Tables 8 through 13 provide some information on this question. When the relationship between the pretest measure and the posttest measure is high and the sample size is large, Table 6 indicates that the variability of the adjustment coefficient is essentially zero. As shown earlier, these are the only conditions under which the procedure estimates the correct effect. Thus, under those conditions the procedure suggested by Kenny is likely to be appropriate. As the sample becomes small and the relationship between measures weakens, the probability of error associated with Kenny's technique increases as does the bias in estimating the effect. These observations were based on the situation when the variance of the measures are equal across time. In situations conforming to the fan spread model, the variability associated with the adjustment coefficient increases and with it the inappropriateness of the gains in standard scores technique. Thus, unless the sample is large and the relationship between the measures is high, the use of gains in standard scores as proposed by Kenny should be avoided. On the other hand, the estimated true score analysis of covariance technique does take the variability of the adjustment coefficient into consideration in determining its standard error. This procedure is, therefore, appropriate when only a single pretest measure is available and data conform to condition 1.

In summary, for the fan spread hypothesis under both condition 1, $\rho = 1$, and condition 2, $\rho \neq 1$, the above findings have indicated that the most desirable analytic strategy of those considered is that of adjusted gains. This approach tested the correct hypothesis under both models of the fan spread condition and with equal to or greater precision than the competing analytic strategies. When only a single pretest was available, estimated true score analysis of covariance was shown to be a more desirable strategy than gains in standard scores for condition 1. This conclusion, however, was limited only to the traditional conceptualization of the fan spread model. This finding, however, was further qualified by suggesting that when (a) the variance of the measures are equal, (b) the relationship between pretest and posttest is high,

and (c) the sample is large; the two procedures estimate the desired effect with equal precision. Finally, when only a single pretest is available and the second model of the fan spread hypothesis is appropriate, no strategy was appropriate. The multiple covariate analysis of covariance as suggested by Keesling and Wiley (1976) was rejected as an inappropriate technique for any condition of the fan spread hypothesis.

A Data Example

To demonstrate the above results, which were derived analytically, a data set was obtained and analyzed. These data were collected on students from seven different elementary schools over a three-year period. Each Spring these students were given the Stanford Achievement Test battery and scores were recorded in the metric of grade equivalents for each subscale of the test as well as total scores. In order to simulate a situation similar to the evaluation of a compensatory education program, random samples of 30 students from the first quartile and 30 students from the second quartile, based on fourth grade total reading scores, were chosen for comparative purposes. Since the students involved were not part of any special program, the only difference between the two groups was their growth rate. The dependent measure chosen to compare the three analytic strategies was the paragraph meaning subtest. Table 14 presents the means in grade equivalents for the hypothetical treatment group (first quartile) and the control group (second quartile) over the three year period. The values in parentheses are the respective standard deviations.

Table 14

Group Means in the Metric of Grade Equivalent Scores and Standard Deviations on the Paragraph Meaning Subtest of the Stanford Achievement Test Battery

	Spring 1973	Spring 1974	Spring 1975
First quartile (treatment)	3.01 (.530)	4.03 (.723)	4.44 (1.218)
Second quartile (control)	3.97 (.453)	4.94 (.939)	5.54 (.939)

These group means are plotted on a time by achievement graph in Figure 7. For the paragraph meaning subtest, students in the first quartile were approximately one grade equivalent behind the students in the second quartile. This difference remained constant over the three-year period. The ratios of the group mean differences to the pooled standard deviations at each point in time were 1.935, 1.089, and 1.012.



Figure 7. Group achievement means plotted across three points in time.

In addition to the three competing analytic strategies considered in this study, traditional analysis of covariance and analysis of variance with residualized gains were also computed. The residualized gains procedure creates a new variable, W, by subtracting from the posttest the product of the regression slope and the pretest measure. The new variable is then used as a dependent variable in the analysis of variance model. The results of these analyses are presented in Table 15. Table 16 presents the effects estimated by each strategy.

These results illustrate several points raised in the earlier discussion. Since there were no differences between the groups except for natural growth rates, the correct conclusion from the analyses should have been that the groups do not differ. The adjusted gain score and gains in standard scores procedures indicated this conclusion. Estimated true score analysis of covariance, on the other hand, indicated a significant difference between the two groups. Examining the relationship between the covariate (Spring 1974 data) and the dependent measure (Spring 1975 data) indicated a correlation of .33. To correct for measurement errors the Kuder-Richardson reliability for internal consistency was used since the more desirable test-retest reliability coefficient was not available. This reliability coefficient equaled .92. The ratio of the correlation coefficient to the reliability coefficient was, therefore, significantly less than unity and little adjustment was made for initial differences.

Table 15

Results of the Data Analyses Using the Gains in Standard Score Strategy, True Score Analysis of Covariance, Adjusted Gain Scores With the Analysis of Variance Model, Adjusted Gain Scores With the Derived Standard Error, Traditional Analysis of Covariance and Analysis of Variance With Residualized Gain Scores

Source	S	d.f.	MS	F	F prob
		Gai	ns in standard scor	res	
Betwee Within	n	1 58	.0835 76.4297	.063	.802
		True sco	ore analysis of cov	ariance	
Betw <mark>ee</mark> Within	n	1 57	4.905 1.065	4.605	.036
	Adjusted	gain scores	(using the analysi	s of variance m	odel)
Betwee Within	n	1 58	.8449 1.2661	.667	.417

Adjusted gain scores (using the derived standard error in a t-test)

$$\frac{\overline{Y}_{p} - \overline{Y}_{c}}{\sqrt{\frac{2}{n} \text{ Var } (\overline{Y}_{p} - \overline{Y}_{c})}} = \frac{-.604 - (-.367)}{\sqrt{\frac{2}{30}} (1.3157)} = \frac{-.237}{.2962} = -.8002$$

$$F = t^{2} = .64037$$

Traditional analysis of covariance

Between Within	1 57	5.690 1.065	5.343	.024		
Analysis of variance with residualized gains						
Between Within	1 58	7.439 1.047	7.108	.010		



The Effect Estimated by the Gains in Standard Scores, True Score Analysis of Covariance, Adjusted Gain Scores, and Traditional Analysis of Covariance

Gains in standard scores:

$$\alpha_{GSS} = \overline{Y}_{p} - \overline{Y}_{c} - \frac{S_{y}}{S_{x}} (\overline{X}_{p} - \overline{X}_{c})$$

= 4.44 - 5.54 - $\frac{1.0785}{.831} (4.03 - 4.94)$
 $\alpha_{GSS} = .08$

Estimated true score analysis of covariance:

$$\alpha_{TS} = \overline{Y}_{p} - \overline{Y}_{c} - \frac{D_{y \cdot x}}{\rho_{xx}} (\overline{X}_{p} - \overline{X}_{c})$$

= 4.44 - 5.54 - $\frac{.44}{.92} (4.03 - 4.94)$
 $\alpha_{TS} = -.665$

Adjusted gain scores:

$$\alpha_{AGS} = Y_p - Y_c - (X_p - X_c) - [(\overline{X}_p - \overline{Z}_p) - (\overline{X}_c - \overline{Z}_c)]$$

= 4.44 - 5.54 - (4.03 - 4.94) - [(4.03 - 3.01) - (4.94 - 3.97)]
$$\alpha_{AGS} = -.24$$

Traditional analysis of covariance and residualized gains:

$$\alpha_{AC} = \overline{Y}_{p} - \overline{Y}_{c} - b_{y \cdot x} (\overline{X}_{p} - \overline{X}_{c})$$

= 4.44 - 5.54 - .44 (4.03 - 4.94)
$$\alpha_{AC} = .70$$

Kenny's procedure indicated no statistically significant differences between the comparison groups when the posttest was the Spring 1975 data and the pretest was the Spring 1974 data. The conclusion would have been different, however, if the period under investigation was the 1973-74 school year. The effect estimated by the strategy would have been the following:

$$\alpha_{GSS} = \overline{Y}_{p} - \overline{Y}_{c} - \frac{Sy}{S_{x}} (\overline{X}_{p} - \overline{X}_{c})$$

$$= 4.03 - 4.94 - \frac{.831}{.492} (3.01 - 3.97)$$

$$= -.91 - 1.689 (-.96)$$

$$= .71,$$

thus indicating a positive treatment effect when there actually was no treatment.

The adjusted gain score approach was computed treating the adjustment as a constant (like Kenny's technique) and as a variable. The degree to which the former analysis is spurious is illustrated by a comparison of the two F-ratios. The F-ratio when the adjustment was treated as a constant was slightly larger, indicating a slightly more powerful test.

The data were also analyzed using the traditional analysis of covariance and the analysis of variance with residualized gains as the dependent variable. The purpose of doing these analyses was to illustrate the inappropriateness of both procedures. The adjustment coefficient used by both procedures is the same, i.e., the observed The average achievement for a particular group was then defined as

Average Achievement =
$$\mu_{\pi}$$
 (t_j - μ_{τ}),

where:

$$\mu_{\pi}$$
 is the average growth rate of the group;
t_j is as defined above; and
 μ_{τ} is the average point in time when the individuals in the
group began to achieve.

Based on this simple model of linear growth, the authors showed that the theoretically correct adjustment coefficient for all linear growth patterns should take the form

$$\frac{\mu_{\pi p} (t_{y} - \mu_{\tau p}) - \mu_{\pi c} (t_{y} - \mu_{\tau c})}{\mu_{\pi p} (t_{x} - \mu_{\tau p}) - \mu_{\pi c} (t_{x} - \mu_{\tau c})},$$

where the subscripts p and c indicate the program and control groups, respectively. This coefficient times the difference between the two groups on the pretest, results in the adjustment factors

$$\mu_{yp} - \mu_{yc} - (\mu_{\pi p} (t_y - \mu_{\tau p}) - \mu_{\pi c} (t_y - \mu_{\tau c})).$$

From this formulation of individual growth, Bryk and Weisberg argued that the gains in standard scores and analysis of covariance with a reliable covariate provided an unbiased estimate of group differences in situations conforming to the traditional fan spread model. Bryk and Weisberg suggested, however, that in situations other than the traditional fan spread, neither of these procedures adequately adjusts for differential linear growth patterns. Their conclusion, therefore, concurs with the findings observed in the present study.

The present study introduced an adjusted gain score strategy. The estimate of group differences determined by this procedure was written as

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - (\mu_{xp} - \mu_{yc}) - [(\mu_{xp} - \mu_{zp}) - (\mu_{xc} - \mu_{zc})].$$

In terms of Bryk and Weisberg's model, this estimate of the group difference can be written as

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - (\mu_{xp} - \mu_{xc}) - \left[\left(\mu_{\pi p} (\mathbf{t}_{x} - \mu_{\tau p}) - \mu_{\pi p} (\mathbf{t}_{z} - \mu_{\tau p}) \right) - \left(\mu_{\pi c} (\mathbf{t}_{x} - \mu_{\tau c}) - \mu_{\pi c} (\mathbf{t}_{z} - \mu_{\tau c}) \right) \right].$$

Where:

$$\mu_{xp} \quad \mu_{\pi p}(t_{x} - \mu_{\tau p});$$

$$\mu_{zp} \quad \mu_{\pi p}(t_{z} - \mu_{\tau p});$$

$$\mu_{xc} \quad \mu_{\pi c}(t_{x} - \mu_{\tau c}); \text{ and }$$

$$\mu_{zc} \quad \mu_{\pi c}(t_{z} - \mu_{\tau c}).$$

Since the average growth rate for each group is assumed to be constant across time, the above estimate of group differences can be simplified to

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - (\mu_{xp} - \mu_{xc}) - \left[\mu_{\pi}[(t_x - \mu_{\tau p}) - (t_z - \mu_{\tau p})] - \mu_{\pi c}[(t_x - \mu_{\tau c}) - (t_z - \mu_{\tau c})]\right].$$

Further simplification is obtained by the following substitution:

$$t_{x} - \mu_{\tau p} = (t_{x} - t_{z}) + (t_{z} - \mu_{\tau p}).$$

Therefore, the adjusted group difference can be written as

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - (\mu_{xp} - \mu_{xc}) - [\mu_{\pi p}(t_x - t_z) - \mu_{\pi c}(t_x - t_z)].$$

If the period of time between the first and second pretests equals the period of intervention, and changing μ_{xp} and μ_{xc} to $\mu_{\pi p}(t_x - \mu_{\tau p})$ and $\mu_{\pi c}(t_x - \mu_{\tau c})$,

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - [\mu_{\pi p}(t_x - \mu_{\tau p}) - \mu_{\pi c}(t_x - \mu_{\tau c}) - \mu_{\pi p}(t_y - t_x) + \mu_{\pi c}(t_y - t_x)].$$

Combining similar terms:

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - \left[\mu_{\pi p} \left[(t_x - \mu_{\tau p}) + (t_y - t_x) \right] - \mu_{\pi c} \left[(t_x - \mu_{\tau c}) + (t_y - t_x) \right] \right].$$

Which can be further simplified to

$$\alpha_{AGS} = \mu_{yp} - \mu_{yc} - [\mu_{\pi p}(t_y - \mu_{\tau p}) - \mu_{\pi c}(t_y - \mu_{\tau c})].$$

This is the adjustment factor identified by Bryk and Weisberg for all linear growth patterns. Thus, the adjusted gain score strategy provides not only the appropriate adjustment for the fan spread model but also any situation in which group growth is linear.

One Group vs. Two Group Research Designs

The adjusted gain score strategy has been shown to be an appropriate technique in situations where groups are changing linearly. Given the assumptions of linear growth and the availability of data from two pretests, the necessity of having an independent control group may be questioned. Under these conditions it might be argued that a treatment group could serve as its own control. This could be achieved by defining a regression line based on the pre-intervention data and predicting posttest scores under a no-treatment effect condition. Differences between the obtained posttest performances and the predicted performances might then be attributed to the treatment. While this approach might be used, it would be difficult to argue that differences between the observed and expected posttest performances were due solely to the treatment. History, maturation, testing, and regression effects are all reasonable threats to internal validity for this approach. The first two threats represent those alternative explanations which can be attributed to changes either outside or within the individual that occur concurrently with the treatment. The last two threats, on the other hand, are factors which could distort the predictive regression line. Testing refers to changes in test performance from the first pretest to the second pretest that are a result of familiarity of test items. Regression effects are those distortions attributed to positive or negative errors of measurement. The use of an independent control group provides a means by which the effects of these threats to internal validity can be controlled as long as those effects are equal across the

two groups. For example, if there are testing effects equaling (a) units on the X measure and this distortion is the same for both groups, then the estimate of the group difference using the adjusted gain score procedure could be written as

$$\mu_{yp} - \mu_{yc} - [(\mu_{xp} + a) - (\mu_{xc} + a)] - [(\mu_{xp} + a - \mu_{zp}) - (\mu_{xc} + a - \mu_{zc})]$$

= $\mu_{yp} - \mu_{yc} - [(\mu_{xp} - \mu_{xc}) + (a - a)] - [(\mu_{xp} - \mu_{zp}) - (\mu_{xc} - \mu_{zc}) + (a - a)].$

Thus, as long as the distortions affect both groups equally, estimates of group differences are still appropriate. The selection by regression or testing interactions refer to the distortions which affect one group to a greater extent than the other group. If this assumption is violated then the group difference estimated is biased.

In addition to the issue of internal validity discussed above, a further consideration is the question of precision. With an independent control group, the contrast of interest is the difference between the adjusted group means. The standard error associated with that contrast can be written as

√2 Var (₩).

The degrees of freedom associated with this test is the sum of the two sample sizes minus 2. In the single group design the contrast of interest is the adjusted group mean minus zero. Its standard error can be written as

Var (W).

The degrees of freedom associated with this test is the number of individuals in the group minus one. Comparing the one group versus the two group design in terms of precision would indicate smaller error variance associated with the former. The two group design, however, has more degrees of freedom. The difference between the degrees of freedom becomes negligible, however, when the sample is large. Thus, the one group design is likely to provide greater precision than the two group design. The problems related to threats to internal validity in the one group design, however, discourage the use of this approach.

Limitations

The adjusted gain score procedure has been presented as an appropriate analytic strategy for situations conforming to the fan spread model. In order to focus attention on what were considered to be the central points for comparison, some assumptions about the circumstances of application have been made. A basic assumption was that it is possible to measure the same individuals repeatedly on the same variable. This may be difficult in a real world setting. The example data examined in the present study to illustrate the analytically derived results, however, has indicated that it is not impossible to obtain such measurements. In schools both the administration and teachers require repeated testing to monitor student progress. These tests, however, may not be appropriate for the adjusted gain score approach since they are unlikely to be the same test or a parallel form of the test. With careful planning, repeated testing of individuals with parallel forms of a test may be possible.

A second assumption has been that there are no selection by regression or selection by testing interactions. Regression effects and testing effects were briefly discussed earlier. If these distortions of the group's growth rate affect both groups equally, then an estimate of group differences is not affected. On the other hand, when there are differential effects associated with these threats to internal validity, then the estimated group difference is biased.

Finally, the adjusted gain score strategy has been based on the assumption that groups grow in a linear fashion. This assumption is likely to be met in situations involving short periods of time. Over extended time periods it seems less likely that a linear model would adequately characterize group changes. The estimated true score analysis of covariance and gains in standard scores strategies also make the same assumption about linear growth. These procedures, however, use data obtained over a shorter period of time than adjusted gain scores and thus less likely to violate the assumption. When the assumption is violated, the adjustment provided by the adjusted gain score strategy can be totally inappropriate. Thus, in situations where the intervention period is extensive, the use of the adjusted gain scores may not be appropriate.

APPENDICES

THE VARIANCE OF THE PRODUCT OF TWO RANDOM VARIABLES

APPENDIX A

APPENDIX A

THE VARIANCE OF THE PRODUCT OF TWO RANDOM VARIABLES

Given: \underline{X} and \underline{Y} stochastically independent such that $E(\underline{X}) = \underline{\xi}, \quad E(\underline{Y}) = \underline{n} \quad \underline{\delta} = \underline{X} - \underline{\xi}$ $Var(\underline{X}) = \Sigma_{\underline{X}}, Var(\underline{Y}) = \Sigma_{\underline{Y}}, \underline{\varepsilon} = \underline{Y} - \underline{\eta}$ $Var[\underline{X'Y}] = Var[\underline{\xi'n} + \underline{\xi'\delta} + \underline{\varepsilon'n} + \underline{\varepsilon'\delta}]$ = Var $(\underline{\xi'n})$ + Var $[\underline{\xi'\delta}]$ + Var $[\underline{\epsilon'n}]$ + Var $[\underline{\epsilon'\delta}]$ + 2 Cov $(\underline{\xi'n}, \underline{\xi'\delta})$ +2 $Cov(\underline{\xi}'\underline{n},\underline{\epsilon}'\underline{n})$ +2 $Cov(\underline{\xi}'\underline{n},\underline{\epsilon}'\underline{\delta})$ +2 $Cov(\underline{\xi}'\underline{\delta},\underline{\epsilon}'\underline{n})$ +2 $Cov(\xi' \delta, \epsilon' \delta)$ +2 $Cov(\epsilon' \eta, \epsilon' \delta)$ $= 0 + \underline{\xi}' \underline{\Sigma}_{\underline{y}} \underline{\xi} + \underline{n}' \underline{\Sigma}_{\underline{x}} \underline{n} + E[(\underline{\varepsilon}' \underline{\delta})^2] + 0 + 0 + 0 + 0 + 2 E[(\underline{\xi}' \underline{\delta})(\underline{\varepsilon}' \underline{\delta})]$ + 2 E [ε'η,ε'δ] $= \underline{\xi}' \Sigma_{\underline{y}} \underline{\xi} + \underline{n}' \Sigma_{\underline{x}} \underline{n} + E [tr \{ (\underline{\varepsilon}' \underline{\delta}) (\underline{\delta}' \varepsilon) \}] + 2 E [tr \{ (\underline{\xi}' \underline{\delta}) (\underline{\delta}' \underline{\varepsilon}) \}]$ + 2 E [tr{($\underline{n}' \varepsilon$)($\underline{\varepsilon}' \delta$)}] $= \underline{\xi}' \Sigma_{\underline{y}} \underline{\xi} + \underline{n}' \Sigma_{\underline{x}} \underline{n} + E [tr \{(\underline{\delta}\underline{\delta}')(\underline{e}\underline{e}')\}] + 2 E [tr \{(\underline{\delta}\underline{\delta}')(\underline{e}\underline{\xi}')\}]$ + 2 E [tr{($\underline{\varepsilon}\underline{\varepsilon}$ ')($\underline{\delta}\underline{\eta}$ ')}] $= \underline{\xi}' \varepsilon_{\underline{y}} \underline{\xi} + \underline{n}' \Sigma_{\underline{x}} \underline{n} + tr\{E(\underline{\delta}\underline{\delta}') E(\underline{\varepsilon}\underline{\varepsilon}')\} + 2 tr\{E(\underline{\delta}\underline{\delta}') E(\underline{\varepsilon}\underline{\xi}')\}$ +2 tr{E($\varepsilon \varepsilon'$) E($\delta \eta'$) $= \underline{\xi}' \Sigma_{y} \underline{\xi} + \underline{\eta}' \varepsilon_{x} \underline{\eta} + tr \{ \Sigma_{y} \Sigma_{x} \}$

If X and Y are scalars rather than vectors:

$$Var(XY) = [E(X)]^{2} \sigma_{y}^{2} + [E(Y)]^{2} \sigma_{x}^{2} + \sigma_{x}^{2} \sigma_{y}^{2}$$

APPENDIX B

THE COVARIANCE OF A RANDOM VARIABLE AND THE PRODUCT OF TWO RANDOM VARIABLES

•

APPENDIX B

THE COVARIANCE OF A RANDOM VARIABLE AND THE PRODUCT OF TWO RANDOM VARIABLES

Given: W = Z - X Y where Z and Y are stochastically independent of X and such that

$$E(X) = \xi, \quad E(Y) = \eta \quad E(Z) = \zeta$$

$$Var(X) = \Sigma_{X}, \quad Var(Y) = \Sigma_{Y}, \quad Var(Z) = \sigma_{Z}^{2}, \quad Cov(Z, \underline{y}) = \sigma_{Z\underline{y}}$$

$$Cov(Z, \underline{X'Y}) = E(Z, \underline{X'Y}) - E(Z) \quad E(\underline{X'Y})$$

$$= E(\underline{X'}) \quad E(Z, \underline{Y}) - \zeta \underline{\xi'\eta}$$

$$= \underline{\xi'}[E(Z, \underline{Y}) - \zeta \underline{\eta}]$$

$$= \underline{\xi'} \quad Cov(Z, \underline{Y})$$

$$= \underline{\xi'}\sigma_{Z\underline{y}}$$

If X and Y are scalars rather than vectors:

= E(X) Cov (Z,Y).

APPENDIX C

THE COVARIANCE OF TWO ADJUSTED VARIABLES

APPENDIX C

THE COVARIANCE OF TWO ADJUSTED VARIABLES

If
$$W_1 = Z_1 - \underline{X}'\underline{Y}_1$$
 and $W_2 = Z_2 - \underline{X}'\underline{Y}_2$

 Z_1 and Y_1 are independent of Z_2 and Y_2 and X independent of Y and ZCov $(W_1W_2) = Cov (\underline{X}'\underline{Y}_1, \underline{X}'\underline{Y}_2)$ $= tr(\Sigma_{\underline{y}_1\underline{y}_2} \Sigma_{\underline{x}}) + n'\Sigma_{\underline{x}}n_2 + \underline{\xi}'\Sigma_{\underline{y}_1\underline{y}_2} \underline{\xi}$ $= \underline{n}_1' \Sigma_{\underline{x}} \underline{n}_2$

If X and Y are scalars rather than vectors

=
$$E(Y_1) E(Y_2) Var(X).$$

APPENDIX D

THE EXPECTED VALUE OF THE RATIO OF TWO NON-INDEPENDENT SAMPLE STANDARD DEVIATIONS SQUARED

de fo Whe If let

APPENDIX D

THE EXPECTED VALUE OF THE RATIO OF TWO NON-INDEPENDENT SAMPLE STANDARD DEVIATIONS SQUARED

The density function of the ratio of two correlated standard deviations was derived by Bose (1935) and Finney (1938) as the following:

$$dF = \frac{2(1-\rho^2)^{\frac{n-1}{2}}}{B(\frac{n-1}{2}, \frac{n-1}{2})} \cdot \frac{W^{n-2}}{(1+W^2)^{n-1}} \left\{ 1 - \frac{4\rho^2 W^2}{(1+W^2)^2} \right\}^{-\frac{n}{2}}$$

Where

$$W = \frac{S_2/\sigma_2}{S_1/\sigma_1}$$

If
$$X^2 - bX + 1 = 0$$

 $X = \frac{1}{2} \left\{ b \pm \sqrt{b^2 - 4} \right\}$ $X^{-1} = \frac{2}{b \pm \sqrt{b^2 - 4}} \cdot \frac{b \pm \sqrt{b^2 - 4}}{b \pm \sqrt{b^2 - 4}} = \frac{b \pm \sqrt{b^2 - 4}}{2}$

let's choose the + sign for x

$$X = \frac{1}{2} \left\{ b + \sqrt{b^2 - 4} \right\} \qquad X^{-1} = \frac{1}{2} \left\{ b - \sqrt{b^2 - 4} \right\}$$
$$X + X^{-1} = b$$
$$X - X^{-1} = \sqrt{b^2 - 4}$$

$$(W + W^{-1})^{2} = W^{2} + W^{-2} + 2 = b + 2 = (1 - \rho^{2})(t + t^{-1})^{2} + 4\rho^{2}$$

$$(W + W^{-1})(W - W^{-1}) \frac{dW}{W} = (1 - \rho^{2})(t + t^{-1})(t - t^{-1}) \frac{dt}{t}$$

$$\left(\frac{W}{1 + W^{2}}\right)^{2} = \left(\frac{1}{W + W^{-1}}\right)^{2} = \frac{1}{(1 - \rho^{2})(t + t^{-1})^{2} + 4\rho^{2}}$$

$$1 - 4\rho^{2} \left(\frac{W}{1 + W^{2}}\right)^{2} = \frac{(1 - \rho^{2})(t + t^{-1})^{2}}{(1 - \rho^{2})(t + t^{-1})^{2} + 4\rho^{2}}$$

$$(W^{2} - W^{2}) = \sqrt{b^{2} - 4} = \sqrt{(b - 2)(b + 2)}$$

$$= \sqrt{\frac{1}{(1 - \rho^{2})(t + t^{-1})^{2} + 4\rho^{2}} \frac{(1 - \rho^{2})(t + t^{-1})^{2} + 4\rho^{2}}{(1 - \rho^{2})(t + t^{-1})^{2} + 4\rho^{2}}$$

$$= \sqrt{(1 - \rho^{2})\frac{(1 - \rho^{2})(t + t^{-1})(t - \rho^{2})(t + t^{-1})^{2} + 4\rho^{2}} \frac{1}{2} \frac{dt}{t}$$

$$= \frac{(1 - \rho^{2})^{\frac{1}{2}}(t + t^{-1})}{(t - t^{-1})(1 - \rho^{2})^{\frac{1}{2}}(t + t^{-1})} \frac{dt}{t}$$

If $X = W^2$, $b = (1 - \rho^2) (t + t^{-1})^2 + 4\rho^2 - 2$

$$\begin{bmatrix} \left(\frac{W}{(1+W^{2})}\right)^{2} \end{bmatrix}^{\frac{n-1}{2}} \left\{ 1 - 4\rho^{2} \left(\frac{W}{1+W^{2}}\right)^{2} \right\}^{-\frac{n}{2}} \frac{dW}{W}$$

$$= \frac{\left\{ (1-\rho^{2})(t+t^{-1})^{2} + 4\rho^{2} \right\}^{\frac{n}{2}} (1-\rho^{2})^{\frac{1}{2}} (t+t^{-1})}{\left\{ (1-\rho^{2})(t+t^{-1})^{2} + 4\rho^{2} \right\}^{\frac{n-1}{2}} (1-\rho^{2})^{\frac{n}{2}} (t+t^{-1})^{n} \left\{ (1-\rho^{2})(t+t^{-1})^{2} + 4\rho^{2} \right\}^{\frac{1}{2}} } \frac{dt}{t}$$

$$= \frac{1}{\left(1-\rho^{2}\right)^{\frac{n-1}{2}} (t+t^{-1})^{n-1}} \frac{dt}{t}$$

$$dF_{t} = \frac{2}{B\left(\frac{n-1}{2}, \frac{n-1}{2}\right)} \frac{1}{(t+t^{-1})^{n-1}} \frac{dt}{t}$$

Let
$$y = \frac{t^2}{1+t^2}$$
 $1 - y = \frac{1}{1+t^2}$ $dy = \frac{2t dt}{(1+t^2)^2}$
 $\frac{2 dt}{(t+t^{-1})^{n-1} t} = \frac{2t^{n-3} t dt}{(1+t^2)^{n-3} (1+t^2)^2} = y^{\frac{n-3}{2}} (1-y)^{\frac{n-3}{2}} dy$
 $dF_y = \frac{y^{\frac{n-3}{2}} (1-y)^{\frac{n-3}{2}} dy}{B(\frac{n-1}{2}, \frac{n-1}{2})}$

Where
$$(W + W^{-1})^2 = \frac{1 - \rho^2}{y(1 - y)} + 4\rho^2$$

 $E(W + W^{-1})^2 = (1 - \rho^2) E \frac{1}{y(1 - y)} + 4\rho^2$
 $E(W + W^{-1}) = E(W^2) + E(W^{-2}) + 2 = 2 \{E(W^2) + 1\}$

$$E \frac{1}{y(1-y)} = \frac{\Gamma(n-1)}{\left[\Gamma\left(\frac{n-1}{2}\right)\right]^2} \int_0^1 y^{\frac{n-5}{2}} (1-y)^{\frac{n-5}{2}} dy$$
$$= \frac{\Gamma(n-1)}{\Gamma(n-3)} \left[\frac{\Gamma(\frac{n-3}{2})}{\Gamma(\frac{n-1}{2})}\right]^2$$
$$= \frac{4(n-2)(n-3)}{(n-3)^2} = \frac{4(n-2)}{n-3}$$
$$E(W^2) = \frac{E(W+W^{-1})^2}{2} - 1 = \frac{(1-\rho^2)}{2(n-3)} + 2\rho^2 - 1$$
$$= \frac{2(n-2)-n+3}{n-3} - \rho^2 \frac{2n-4-2n+6}{n-3} = \frac{(n-1)-2\rho^2}{n-3}$$
$$W^2 = \frac{\sigma_1^2}{\sigma_2^2} \frac{S_2^2}{S_1^2}$$

- 1

$$E\left(\frac{S_{2}^{2}}{S_{1}^{2}}\right) = \frac{(n-1)-2\rho^{2}}{n-1} \frac{\sigma_{2}^{2}}{\sigma_{1}^{2}}.$$

APPENDIX E

THE EXPECTED VALUE OF THE RATIO OF TWO NON-INDEPENDENT SAMPLE STANDARD DEVIATIONS
APPENDIX E

THE EXPECTED VALUE OF THE RATIO OF TWO NON-INDEPENDENT SAMPLE STANDARD DEVIATIONS

From Appendix D:

$$W + W^{-1} = \sqrt{\frac{1 - \rho^{2}}{y(1 - y)} + 4\rho^{2}}$$

$$E(W + W^{-1}) = 2E(W)$$

$$E(W) = \frac{1}{2} \int_{0}^{1} \sqrt{\frac{1 - \rho^{2}}{y(1 - y)} + 4\rho^{2}} \frac{\left[y(1 - y)\right]^{\frac{n-3}{2}}}{B\left(\frac{n-1}{2}, \frac{n-1}{2}\right)} dy$$

$$= \frac{\Gamma(n-1)}{\left[\Gamma\left(\frac{n-1}{2}\right)\right]^{2}} \int_{0}^{\cdot 5} \sqrt{1 - \rho^{2} + 4\rho^{2}} y(1 - y) \left[y(1 - y)\right]^{\frac{n-4}{2}} dy.$$

The above integral was evaluated with the DCADRE program on the CDC6500 computer.

$$W = \frac{\sigma_1}{\sigma_2} \frac{S_2}{S_1}$$

$$E\left(\frac{S_{2}}{S_{1}}\right) = \frac{\Gamma(n-1)}{\left[\Gamma\left(\frac{n-1}{2}\right)\right]^{2}} \int_{0}^{.5} \sqrt{1-\rho^{2}+4\rho^{2} y(1-y)} \left[y(1-y)\right]^{\frac{n-4}{2}} dy \frac{\sigma_{2}}{\sigma_{1}}.$$

APPENDIX F

THE EXPECTED VALUE OF A SAMPLE REGRESSION COEFFICIENT

APPENDIX F

THE EXPECTED VALUE OF A SAMPLE REGRESSION COEFFICIENT

Density function for $(\beta_{y*x}) =$

$$dF = \frac{\Gamma\left(\frac{n}{2}\right)(1-\rho^{2})^{\frac{n-1}{2}}}{\sqrt{\pi} \Gamma\left(\frac{n-1}{2}\right)} \left(\frac{\sigma_{y}}{\sigma_{x}}\right)^{n-1} \frac{db}{\left\{\frac{\sigma_{y}^{2}}{\sigma_{x}^{2}} \left(1-\rho^{2}\right) + \left(b-\frac{\rho \sigma_{y}}{\sigma_{x}}\right)^{2}\right\}^{n/2}}$$

$$\operatorname{let} \frac{\sigma_{\mathbf{y}}}{\sigma_{\mathbf{x}}} = k$$

$$E(b_{\mathbf{y}\cdot\mathbf{x}}) = \frac{\Gamma(\frac{n}{2})(1-\rho^{2})^{\frac{n-1}{2}}}{\sqrt{\pi} \Gamma(\frac{n-1}{2})} k^{n-1} \int_{-\infty}^{+\infty} \frac{b \, db}{\left\{k^{2}(1-\rho^{2})+(b-\rho k)^{2}\right\}^{n/2}}$$

$$=\frac{\Gamma\left(\frac{n}{2}\right)(1-\rho^{2})^{\frac{n-1}{2}}k^{n-1}}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{b\ db}{\left\{\left(1+\frac{(b-\rho k)^{2}}{k^{2}(1-\rho^{2})}\right)\left(k^{2}(1-\rho^{2})\right)\right\}^{n/2}}$$

$$= \frac{\Gamma\left(\frac{n}{2}\right)(1-\rho^{2})^{\frac{n-1}{2}}k^{n-1}}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)(1-\rho^{2})^{n/2}}k^{n}\int_{-\infty}^{+\infty} \frac{b \ db}{\left\{1+\frac{(b-\rho k)^{2}}{k^{2}(1-\rho^{2})}\right\}^{n/2}}$$

-

let
$$t = b - \rho k$$

 $b = t + \rho k$
 $db = dt$
 $v = k^2(1 - \rho^2)$

$$\int_{-\infty}^{+\infty} \frac{t + \rho k \, dt}{\left|1 - \frac{t^2}{\nu}\right|^{n/2}}$$

$$\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{n-1}{2}\right) \kappa (1-\rho^2)^{1/2}} \left[\int_{-\infty}^{+\infty} \frac{t \, dt}{\left(1-\frac{t^2}{\nu}\right)^{n-2}} + \rho k \int_{-\infty}^{+\infty} \frac{1 \, dt}{\left(1+\frac{t^2}{\nu}\right)^{n-2}} \right]$$

$$let \quad y = \frac{1}{1+\frac{t^2}{\nu}} = \frac{1}{\frac{\nu+t^2}{\nu}} = -\frac{\nu}{\nu+t^2}$$

$$\frac{\nu}{\nu} = \nu + t^2$$

$$\nu\left(\frac{1}{y}-1\right) = t^2$$

$$\nu\left(\frac{1-y}{y}\right) = t^2$$

$$t = \frac{(\nu-\nu y)^{1/2}}{y^{1/2}}$$

$$dt = \frac{y^{1/2} \left[\frac{1}{2} (v - vy)^{1/2} (-v) \right] - \left[(v - vy)^{1/2} \frac{1}{2} (y) \right]^{1/2}}{y}$$
$$= \frac{-\frac{1}{2} \left[\frac{y^{1/2} v}{[v - (1 - y)]^{1/2}} + \frac{v^{1/2} (1 - y)^{1/2}}{y^{1/2}} \right]}{y}$$
$$= \frac{-\frac{1}{2} \left[\frac{y^{1/2} y^{1/2} v^{1/2} + v^{1/2} (1 - y)^{1/2} (1 - y)^{1/2}}{[(1 - y) y]^{1/2}} \right]}{y}$$

$$= \frac{\frac{-v^{1/2}}{2} \left[\frac{y+1-y}{[(1-y) y]^{1/2}} \right]}{y}$$

$$dt = \frac{-v^{1/2}}{2} \frac{1}{[(1 - y) y]^{1/2} y} dy$$

$$k(1 - \rho^2)^{1/2} = \sqrt{\nu}$$

$$\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)}\int_{-\infty}^{+\infty}\frac{t\,dt}{\left(1+\frac{t^{2}}{\nu}\right)^{n/2}\nu^{1/2}}=\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)}\left[\int_{-\infty}^{0}\frac{t\,dt}{\left(1+\frac{t^{2}}{\nu}\right)^{n-2}\nu^{1/2}}\right]$$

$$+ \int_0^\infty \frac{t \, dt}{\left(1 + \frac{t^2}{\nu}\right)^{n/2} \nu^{1/2}}$$

$$\int_{-\infty}^{0} \frac{t \, dt}{\left(1 + \frac{t^2}{\nu}\right)^{n/2} \nu^{1/2}}$$

$$= \int_{0}^{1} y^{\frac{n}{2}} \left[v^{1/2} \frac{(1-y)^{1/2}}{y^{1/2}} \right] \left[\frac{-v^{1/2}}{2} \frac{1}{\left[(1-y)^{1/2} y^{1/2} \right] y v^{1/2}} \right]$$

.

$$= \frac{-v^{1/2}}{2} \int_{0}^{1} y^{\frac{n-4}{2}} dy = 0.$$

Similarly,
$$\int_{0}^{\infty} \frac{t \, dt}{\left(1 + \frac{t^2}{\nu}\right)^{n/2} \nu^{1/2}} = 0$$

$$\int_{-\infty}^{+\infty} \frac{1}{\left(1 + \frac{t^2}{\nu}\right)^{n/2} \nu^{1/2}} = \int_{-\infty}^{0} \frac{1}{\left(1 + \frac{t^2}{\nu}\right)^{n/2} \nu^{1/2}} + \int_{0}^{\infty} \frac{1}{\left(1 + \frac{t^2}{\nu}\right)^{n/2} \nu^{1/2}} = -2 \int_{0}^{1} y^{\frac{n}{2}} \left[-\frac{\nu^{1/2}}{2} \frac{1}{(1 - y)^{1/2} y^{3/2} \nu^{1/2}} \right] \\ = \frac{\Gamma\left(\frac{n - 1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} .$$

Substituting the two integrals

$$E(b_{y\cdot x}) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{n-1}{2}\right)} \left[0 + \rho k \left(\frac{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \right) \right]$$

$$E(b_{y \cdot x}) = \rho k$$
$$= \rho \frac{\sigma_y}{\sigma_x}.$$

APPENDIX G

.

THE EXPECTED VALUE OF A SAMPLE REGRESSION COEFFICIENT SQUARED

APPENDIX G

THE EXPECTED VALUE OF A SAMPLE REGRESSION COEFFICIENT SQUARED

$$dF = \frac{\Gamma\left(\frac{n}{2}\right)(1-\rho^2)^{\frac{n-1}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)} \left(\frac{\sigma_y}{\sigma_x}\right)^{n-1} \frac{db}{\left\{\frac{\sigma_y^2}{\sigma_x^2}\left(1-\rho^2\right) + \left(b-\frac{\rho\sigma_y}{\sigma_x}\right)^2\right\}^{n/2}}$$

let $\frac{\sigma y}{\sigma x} = k$

$$E(b^{2}) = \frac{\Gamma(\frac{n}{2})(1-\rho^{2})^{\frac{n-1}{2}}}{\sqrt{\pi}\Gamma(\frac{n-1}{2})} k^{n-1} \int_{-\infty}^{+\infty} \frac{b^{2} db}{\left\{k^{2}(1-\rho^{2})+(b-\rho k)^{2}\right\}^{n/2}}$$

$$\int_{-\infty}^{+\infty} \frac{b^2 db}{\left| \left(1 + \frac{(b-\rho k)^2}{k^2 (1-\rho^2)} \right) \left(k^2 (1-\rho^2) \right) \right|^{n/2}}$$

$$\frac{\Gamma\left(\frac{n}{2}\right)(1-\rho^{2})^{\frac{n-1}{2}}k^{n-1}}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)(1-\rho^{2})^{n/2}k^{n}}\int_{-\infty}^{+\infty}\frac{b^{2}}{\left\{1+\frac{(b-\rho k)^{2}}{k^{2}(1-\rho^{2})}\right\}^{n/2}}$$

$$t = b - \rho k$$

$$b = t + \rho k$$

$$db = dt$$

$$v = k^2 (1 - \rho^2)$$

$$\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)k\,\left(1-\rho^2\right)^{1/2}}\left[\int_{-\infty}^{+\infty}\frac{t^2+2t\rho k+\left(\rho k\right)^2}{\left|1+\frac{t^2}{\nu}\right|^{n/2}}\,dt\right]$$

$$\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n-1}{2}\right)} \left[\int_{-\infty}^{+\infty} \frac{t^2}{\left\{1+\frac{t^2}{\nu}\right\}^{n/2} \nu^{1/2}} dt \right]$$

+
$$2pk \int_{-\infty}^{+\infty} \frac{t dt}{\left(1 + \frac{t^2}{\nu}\right)^{n/2} \nu^{1/2}} + (\rho k)^2 \int_{-\infty}^{+\infty} \frac{1 dt}{\left(1 + \frac{t^2}{\nu}\right)^{n/2} \nu^{1/2}}$$

$$\int_{-\infty}^{+\infty} \frac{t^2}{\left\{1 + \frac{t^2}{\nu}\right\}^{n/2}} \frac{dt}{\nu^{1/2}} = -2 \int_{0}^{+\infty} \frac{t^2}{\left\{1 + \frac{t^2}{\nu}\right\}^{n/2}}$$

$$y = \frac{1}{1 + \frac{t^{2}}{v}}$$

$$t^{2} = v\left(\frac{1-y}{y}\right)$$

$$dt = -\frac{v^{1/2}}{2} \left[\frac{1}{((1-y) \ y)^{1/2} \ y}\right] dy$$

$$= v^{1/2} \int_{0}^{+\infty} y^{n/2} \left[v\left(\frac{1-y}{y}\right)\right] \frac{1}{(1-y)^{1/2} \ y^{3/2}} \frac{dy}{v^{1/2}}$$

$$= v \int_{0}^{1} y^{\frac{n-5}{2}} (1-y)^{1/2} \ dy$$

$$= \frac{v \Gamma\left(\frac{n-3}{2}\right) \Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}$$

From Appendix F

$$\int_{-\infty}^{+\infty} \frac{t \, dt}{\left\{1 - \frac{t^2}{\nu}\right\}^{n/2}} = 0 \quad \text{and} \quad \int_{\infty}^{+\infty} \frac{1 \, dt}{\left\{1 + \frac{t^2}{\nu}\right\}^{n/2}} = \frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}$$
$$= \left(b_{y \cdot x}^2\right) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi} \Gamma \frac{n-1}{2}} \left[\frac{\nu \Gamma\left(\frac{n-3}{2}\right)\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} + 2\rho k(0) + (\rho k)^2 \frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma \frac{1}{2}}{\Gamma\left(\frac{n}{2}\right)}\right]$$

$$= \frac{v}{n-3} + \rho^2 k^2$$

$$= \frac{(1-\rho^2)\frac{\sigma_y^2}{\sigma_x^2}}{n-3} + \rho^2 \frac{\sigma_y^2}{\sigma_x^2} .$$

APPENDIX H

THE VARIANCE OF THE ADJUSTED GAIN SCORE VARIABLE

APPENDIX H

THE VARIANCE OF THE ADJUSTED GAIN SCORE VARIABLE

- Let: Z represent the first pretest
 - X represent the pretest administered immediately prior to the intervention
 - Y represent the post-intervention test performance

Then the adjusted score is determined as follows:

$$W = Y - X - [\overline{X} - \overline{Z}]$$
$$W = Y - X - \overline{X} + \overline{Z}$$

$$Var (W) = Var [Y - X - \overline{X} + \overline{Z}]$$

$$= Var (Y) + Var (X) + Var (\overline{X}) + Var (\overline{Z})$$

$$- 2Cov(Y,X) - 2Cov(Y,\overline{X}) + 2Cov(Y,\overline{Z}) + 2Cov(X,\overline{X}) - 2Cov(X,\overline{Z}) - 2Cov(\overline{X},\overline{Z})$$

$$Cov(Y,\overline{X}) = \frac{\sum_{\Sigma} n (Y - \mu_{Y})(\overline{X} - \mu_{X})}{Nn}$$

$$= \sum_{\Sigma} n \overline{Y} \overline{X} - \sum_{\Sigma} \sum_{\Sigma} Y \mu_{X} - \sum_{\Sigma} \sum_{\Sigma} \overline{X} \mu_{Y} + \sum_{\Sigma} \sum_{\mu_{Y}} \mu_{X}/Nn$$

$$= \sum_{\Sigma} n \overline{Y} \overline{X} - \sum_{\Sigma} n \overline{Y} \mu_{X} - Nn \mu_{X}\mu_{Y} + Nn \mu_{Y}\mu_{X}/Nn$$

$$= \sum_{\Sigma} \overline{Y} \overline{X} - \mu_{X} \sum_{\overline{Y}} \overline{Y} /N$$

$$= \sum_{\Sigma} \overline{Y} \overline{X} - \mu_{X} \sum_{\overline{Y}} \overline{Y} /N$$

$$= Cov (\overline{Y}, \overline{X})$$

Cov $(Y,\overline{Z}) = Cov (\overline{Y},\overline{Z})$ Cov $(X,\overline{X}) = Var (\overline{X})$ Cov $(X,\overline{Z}) = Cov (\overline{X},\overline{Z})$

Assume a single population with observations on two variables X and Y. Select N random samples of size n from that population, then:

Cov
$$(\overline{X}, \overline{Y}) = \frac{Cov (X, Y)}{n}$$

Cov $(\overline{X}, \overline{Y}) = \frac{\sum_{n=1}^{N} (\overline{X} - \mu_{X})(\overline{Y} - \mu_{y})}{N}$
 $= \sum_{n=1}^{N} (\overline{X} \ \overline{Y} - \overline{X} \ \mu_{y} - \overline{Y} \ \mu_{x} + \mu_{x} \ \mu_{y})/N$
 $= \left(\sum_{n=1}^{N} \overline{X} \ \overline{Y} - \sum_{n=1}^{N} \overline{X} \ \mu_{y} - \sum_{n=1}^{N} \overline{Y} \ \mu_{x} + \sum_{n=1}^{N} \mu_{x} \ \mu_{y}\right)/N$
 $= \left(\sum_{n=1}^{N} \overline{X} \ \overline{Y} - N \ \mu_{x} \ \mu_{y} - N \ \mu_{y} \ \mu_{x} + N \ \mu_{x} \ \mu_{y}\right)/N$
 $= \left(\sum_{n=1}^{N} \overline{X} \ \overline{Y} - N \ \mu_{x} \ \mu_{y}\right)/N$
 $= \left(\sum_{n=1}^{N} \overline{X} \ \overline{Y} - N \ \mu_{x} \ \mu_{y}\right)/N$
 $\sum_{n=1}^{N} \overline{X} \ \overline{Y} = \sum_{n=1}^{N} \frac{\sum_{n=1}^{N} \sum_{n=1}^{N} \mu_{n}}{n^{2}}$
 $\sum_{n=1}^{N} x \ \overline{Y} = \sum_{n=1}^{N} x_{i}Y_{i} + n(n-1) \ \mu_{x} \ \mu_{y}$
i.e., $(X_{1} + X_{2})(Y_{1} + Y_{2}) = X_{1}Y_{1} + X_{2}Y_{2} + X_{1}Y_{2} + X_{2}Y$

assuming that all pairs where $i \neq i'$ are independent and there are n(n-1) such pairs.

$$\frac{1}{n}\sum_{\Sigma}^{N} \begin{pmatrix} n \\ \Sigma X \\ \Sigma & \Sigma \end{pmatrix} = \frac{1}{n^{2}} \begin{bmatrix} N \\ \Sigma \\ \Sigma \\ \Sigma \\ \Sigma \\ X \end{pmatrix} + \frac{N}{2} \begin{bmatrix} n^{2} \\ n^{2} \\ \mu_{X} \\ \mu_{Y} \end{bmatrix} + \frac{1}{n^{2}}\sum_{\Sigma}^{N} n^{2} \\ \mu_{X} \\ \mu_{Y} \\ \mu_{Y} \end{bmatrix} / N$$

$$\frac{1}{n} \begin{bmatrix} N \\ \Sigma \\ \Sigma \\ \Sigma \\ X \\ Y - N \\ n \\ \mu_{X} \\ \mu_{Y} \end{bmatrix} / N n$$

$$\frac{1}{n} \begin{bmatrix} N \\ \Sigma \\ \Sigma \\ X \\ Y - N \\ n \\ \mu_{X} \\ \mu_{Y} \end{bmatrix} / N n$$

$$= \frac{1}{n} \begin{bmatrix} N \\ \Sigma \\ \Sigma \\ X \\ Y - N \\ n \\ \mu_{X} \\ \mu_{Y} \end{bmatrix} / N n$$

$$= \frac{1}{n} \begin{bmatrix} N \\ \Sigma \\ \Sigma \\ X \\ Y - N \\ n \\ \mu_{X} \\ \mu_{Y} \end{bmatrix} / N n$$

 $Var (W) = Var(Y) + Var(X) + Var(\overline{X}) + Var(\overline{Z}) - 2Cov(Y,X) - 2Cov(\overline{Y},\overline{X}) + 2Cov(\overline{Y},\overline{Z})$ $+ 2Var(\overline{X}) - 2Cov(\overline{X},\overline{Z}) - 2Cov(\overline{X},\overline{Z})$ $Assuming independent samples <math>Cov(\overline{X},\overline{Y}) = \frac{Cov(X,Y)}{n}$.

$$\begin{aligned} \operatorname{Var}(W) &= \operatorname{Var}(Y) + \operatorname{Var}(X) + \frac{1}{n} \operatorname{Var}(X) + \frac{1}{n} \operatorname{Var}(Z) - 2\operatorname{Cov}(Y,X) - \frac{2}{n} \operatorname{Cov}(Y,X) \\ &+ \frac{2}{n} \operatorname{Cov}(Y,Z) + \frac{2}{n} \operatorname{Var}(X) - \frac{4}{n} \operatorname{Cov}(X,Z) \\ &= \operatorname{Var}(Y) + \left[\operatorname{Var}(X) + \frac{1}{n} \operatorname{Var}(X) + \frac{2}{n} \operatorname{Var}(X) \right] + \left[- 2 \operatorname{Cov}(X,Y) - \frac{2}{n} \operatorname{Cov}(X,Y) \right] \\ &+ \frac{1}{n} \operatorname{Var}(Z) + \frac{2}{n} \operatorname{Cov}(Y,Z) - \frac{4}{n} \operatorname{Cov}(X,Z) \\ &= \operatorname{Var}(Y) + \left(1 + \frac{3}{n} \right) \operatorname{Var}(X) - \left(2 + \frac{2}{n} \right) \operatorname{Cov}(X,Y) \\ &+ \frac{1}{n} \left[\operatorname{Var}(Z) + 2\operatorname{Cov}(Y,Z) - 4\operatorname{Cov}(X,Z) - 4\operatorname{Cov}(X,Z) \right] \\ \end{aligned}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

- Baughman, E. E. <u>Negro and white children: A psychological study in</u> the rural South. New York: Academic Press, 1968.
- Bereiter, C. Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), <u>Problems in measuring change</u>. Madison: University of Wisconsin Press, 1963.
- Berkson, J. Are there two regressions? <u>Journal of American</u> <u>Statistical Association</u>, 1950, 45, 164-180.
- Bose, S. S. On the distribution of the ratio of variances of two samples drawn from a given bivariate correlated population. Sankhya, 1935, <u>2</u>, 65.
- Bryk, A. S., & Weisberg, H. Use of the non-equivalent control group design when subjects are growing. To appear in <u>Psychological</u> <u>Bulletin</u>, 1977.
- Campbell, D. T., & Stanley, J. C. <u>Experimental and quasi-experimental</u> <u>designs for research</u>. Chicago: Rand McNally, 1966.
- Campbell, D. T., & Boruch, R. F. Making the case for randomized assignment to treatments by considering the alternatives: Six ways in which quasi-experimental evaluations in compensatory education tend to underestimate effects. In C. A. Bennett (Ed.), Evaluation and experiment. New York: Academic Press, 1975.
- Campbell, D. T., & Erlebacher, A. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Hellmuth (Ed.), <u>Compensatory</u> <u>education: A national debate</u> (Vol. 3). <u>The disadvantaged child</u>. New York: Brunner/Mazel, 1970.
- Campbell, D. T. Temporal changes in treatment-effect correlations: A quasi-experimental model for institutional records and longitudinal studies. In G. V. Glass (Ed.), <u>Proceedings of</u> <u>the 1970 invitational conference on testing problems: the</u> <u>promise and perils of educational information systems</u>. New York: Educational Testing Service, 1971.
- Cronbach, L. J., & Furby, L. How we should measure 'change' or should we? <u>Psychological Bulletin</u>, 1970, <u>74</u>, 68-80.

- de Boor, C. CADRE: An algorithm for numerical quadrature. In
 J. Rice (Ed.), Mathematical software. New York: Academic
 Press, 1971.
- De Gracie, J. Analysis of covariance when the concomitant variable is measured with error. Unpublished Ph.D. dissertation, Iowa State University, 1968.
- Elashoff, J. D. Analysis of covariance: A delicate instrument. American Educational Research Journal, 1969, 6, 383-401.
- Finney, D. J. The distribution of the ratio of estimates of the two variances in a sample from a normal bivariate population. Biometrika, 1938, 10, 190-193.
- Fennessey, J. <u>Incentives in education project: Impact evaluation</u> <u>report</u> (Final report under contract no. OEC-0-71-4770 to U.S. Office of Education). Washington, D.C.: The Planar Corporation, 1972.
- Harnquist, K. Relative changes in intelligence from 13 to 18. Scandinavian Journal of Psychology, 1968, 9, 50-82.
- Harnquist, K. Correcting analysis of covariance for unreliability of the control variable. Unpublished paper, University of Gothenburg, Sweden, 1958.
- Jöreskog, K. G., & van Thillo, M. LISREL--A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables. <u>RB-72-56</u>. Princeton, N.J.: Educational Testing Service, 1972.
- Jöreskog, K. G. A general method for estimating a linear structural equation system. In A. S. Goldberger and Duncan (Eds.), <u>Structural equations models in the social sciences</u>. New York: Seminar Press, 1973.
- Keesling, J. W., & Wiley, D. Measurement error and the analysis of quasi-experimental data, version IV. Unpublished manuscript, 1976.
- Kendall, M. G., & Stuart, A. <u>The advanced theory of statistics</u> (Vol. 1). <u>Distribution theory</u>. New York: Hafner Publishing Co., 1952.
- Kenny, D. A quasi-experimental approach to assessing treatment effects in the non-equivalent control group design. <u>Psychological Bulletin</u>, 1975, 82, 345-362.

- Linn, R., & Slinde, J. The determination of the significance of change between pre and posttesting periods. <u>Review of Educational</u> <u>Research</u>, 1977, <u>47</u>, 121-150.
- Lord, F. M. The measurement of growth. <u>Educational and Psychological</u> <u>Measurement</u>, 1956, 16, 421-437.
- Lord, F. M. Further problems in the measurement of growth. Educational and Psychological Measurement, 1958, 18, 437-451.
- Lord, F. M. Large-sample covariance analysis when the control variable is fallible. <u>Journal of the American Statistical Association</u>, 1960, <u>55</u>, 307-321.
- Lord, F. M. Elementary models for measuring change. In C. W. Harris (Ed.), <u>Problems in measuring change</u>. Madison: University of Wisconsin Press, 1963.
- Lord, F. M. A paradox of interpretation of group comparisons. Psychological Bulletin, 1967, 68, 304-305.
- Lord, F. M. Statistical adjustments when comparing pre-existing groups. <u>Psychological Bulletin</u>, 1969, 72, 336-337.
- Marks, E., & Martin, C. G. Further comments relating to the measurement of change. <u>American Educational Research Journal</u>, 1973, <u>10</u>, 179-191.
- McNemar, Q. On growth measurement. <u>Educational and Psychological</u> Measurement, 1958, 18, 47-55.
- Osborne, R. T., & Suddick, D. E. A longitudinal investigation of the intellectual differentiation hypothesis. Journal of Genetic <u>Psychology</u>, 1972, 121, 83-89.
- Porter, A. C. The effects of using fallible variables in the analysis of covariance. Unpublished Ph.D. dissertation, Madison: University of Wisconsin, 1967.
- Porter, A. C. How errors of measurement affect ANOVA, regression analyses, ANCOVA and factor analyses. Paper presented at the AERA Convention, New York, 1971.
- Porter, A. C. Analysis strategies for some common evaluation paradigms. Paper presented at the AERA Convention, New Orleans, February 1973.
- Porter, A. C., & Chibucos, T. R. Selecting an analysis strategy. In G. Borich (Ed.), <u>Evaluating Educational Programs and Products</u>. Educational Technology Press, 1974.

- Pravalpruk, K. Investigation of two proposed solutions to the multiple fallible covariable problem for quasi-experiments. Unpublished Ph.D. dissertation, East Lansing: Michigan State University, 1974.
- Richards, J. M., Jr. A simulation study of the use of change measures to compare educational programs. <u>American Educational Research</u> <u>Journal</u>, 1975, 12, 299-311.
- Stroud, T. W. Comparing conditional means and variances in a regression model with measurement errors of known variances. <u>Journal of the American Statistical Association</u>, 1972, <u>67</u>, 407-412.
- Stroud, T. W. Comparing regressions when measurement error variances are known. Psychometrika, 1974, 39, 53-68.

