DOCUMENT RESUME

ED 342 791 TM 017 948

AUTHOR Bump, Wren M.
TITLE The Homogeneity of Regression Assumption in the Analysis of Covariance.
PUB DATE Feb 92
NOTE 17p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Houston, TX, January 31-February 2, 1992).
PUB TYPE Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Analysis of Covariance; Equations (Mathematics); Error of Measurement; *Experimental Groups; Heuristics; Mathematical Models; *Regression (Statistics); *Research Design; *Sampling
IDENTIFIERS *Homogeneity of Regression; Power (Statistics)

ABSTRACT

An analysis of covariance (ANCOVA) is done to correct for chance differences that occur when subjects are assigned randomly to treatment groups. When properly used, this correction results in adjustment of the group means for pre-existing differences caused by sampling error and reduction of the size of the error variance of the analysis. The adjustment of the means is done to reduce bias that may be caused by the differences. This hoped-for increase in power is a major advantage of ANCOVA. However, the inappropriate use of ANCOVA appears to be the rule rather than the exception. This paper explains the homogeneity of regression assumption and why it is so important to evaluate this assumption before conducting an ANCOVA. Small heuristic data sets (3 groups of 12 entries each for intelligence quotient and achievement) are used to make the discussion concrete. Three tables present the data sets, and five figures illustrate their application. A seven-item list of references is included. (SLD)

The Homogeneity of Regression Assumption in the Analysis of Covariance

Wren M. Bump
Texas A&M University 77843-3242

Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX, February 1, 1992.

Abstract

The inappropriate use of ANCOVA appears to be the rule rather than the exception (Thompson, 1991). As Keppel and Zedeck (1989, e.g., pp. 455, 456, 466, 478-479, 480) repeatedly and emphatically argue, ANCOVA is only appropriate for use in conjunction with randomly assigned groups. Keppel and Zedeck (1989, p. 481) cogently explain why and point out the importance of the homogeneity of regression assumption. As they note:

> It is somewhat depressing to note that while all statistical methodology books continue to stress [would that this were true] the conclusion that ANCOVA should not be used in quasi-experimental designs, misapplications of the procedure are still committed and reported in the literature. (p. 482)

Anyone contemplating an ANCOVA should seriously consider the trenchant arguments made by Campbell and Erlebacher (1975). The purpose of this paper is to explain the homogeneity of regression assumption, and why it is so important to evaluate this assumption before conducting an analysis of covariance. Small heuristic data sets will be employed to make this discussion concrete.

The Homogeneity of Regression Assumption in the Analysis of Covariance

An analysis of covariance (ANCOVA) is done to correct for chance differences that occur when subjects are assigned **randomly** to the treatment groups. When properly used this correction results in (a) an **adjustment of the group means** for those pre-existing differences caused by sampling error and (b) the size of the error variance in the analysis is reduced, thus increasing statistical power. The adjustment of the means is done to reduce bias that may be caused by the differences. This hoped-for increase in power, according to Huitema (1980, p. 13), is "the major payoff in selecting the analysis of covariance".

Correct Usage of ANCOVA

Before we look at a graphical and mathematical representation of ANCOVA, let's look at a simplified "pie" version of ANCOVA. The "pie" version means that the sum of squares total is represented by a pie. We want to eat (explain) as much of the pie as we can. (In Figure 1, it is the Y "pie" that we want to eat.) The more Y pie we eat, the more variance we can explain. Any uneaten (or unexplained) part is error. The eaten part of the pie is the percent of overlap of X and Y and the covariate (see Figure 1). This percent of overlap represents r^2 which has several names - coefficient of determination, common variance and effect size.

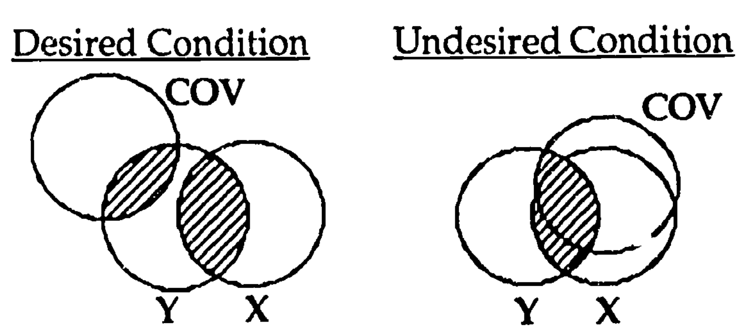

Figure 1 Conditions showing when ANCOVA use would be desired and when ANCOVA use would be undesired

Loftin and Madison (1991, p. 134) list five conditions that must be met for correct ANCOVA usage. The first two of these conditions will be examined and illustrated. Referring to Figure 1 we will see why these conditions must be met. The first condition is that the covariate should be an independent variable (which means it is pre-existing) that is highly correlated with the variable. "If the correlation is not high, the covariate will do little to reduce the error sum of squares, and this is the primary objective of ANCOVA" (Loftin & Madison,1991, p. 135). The desired condition, in Figure 1, shows that the covariate is highly correlated with Y, that is, it eats a good portion of the Y pie.

The second condition is that the covariate should be uncorrelated with the independent variable. The desired condition, in Figure 1, shows that the covariate and the independent variable X are uncorrelated, that is, they eat different parts of the Y pie. The covariate adjustment is always made first and all the error sum of square will be attributed to the the covariate if there is a correlation between the covariate and independent variable (see Figure 1, undesired condition). In this case, the covariate and the independent variable are highly correlated and are eating much of the same parts of the pie. The problem with this is that the covariate will get the first credit for eating the pie and X will be credited only with the "leftovers" it eats.

ANCOVA as an integration of ANOVA and ANOVAR

Huitema introduces ANCOVA as an integration of the analysis of variance (ANOVA) and "ANOVA of regression" (ANOVAR). The ANOVA model accounts for between-group variance and the ANOVAR model accounts for regression variance. "The ANCOVA model treats *both* between-group and regression variance as systematic (nonerror) components" (Huitema, 1980, p. 25).

Figure 2 presents a comparison of the three analyses, both mathematically and graphically. When the ANOVAR model is added to the ANOVA, the result is the ANCOVA model. The power that Huitema refers to comes from the reduction of the error term. Huitema (1980) states:

> If the assumptions for the analysis of covariance are met for a given set of data, the error term will be smaller than in ANOVA because much of the within-group variability will be accounted for and removed by the regression of the dependent variable on the covariate. Since the result of the smaller error term is an increase in power, it is quite possible that data analyzed by using ANCOVA will yield highly significant results where ANOVA yields nonsignificant results. (p. 25)

<u>**Technique**</u>	<u>**Model**</u>
ANOVA	$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$
ANOVA of Regression	$Y_i = \beta_0 + \beta_1 X + \varepsilon_i$ or $y_i = \beta_1 (X_i - \overline{X}..) + \varepsilon_i$
ANCOVA	$Y_{ij} = \mu + \alpha_j + \beta_1 (X_{ij} - X..) + \varepsilon_{ij}$

<u>**Variance Partition**</u>

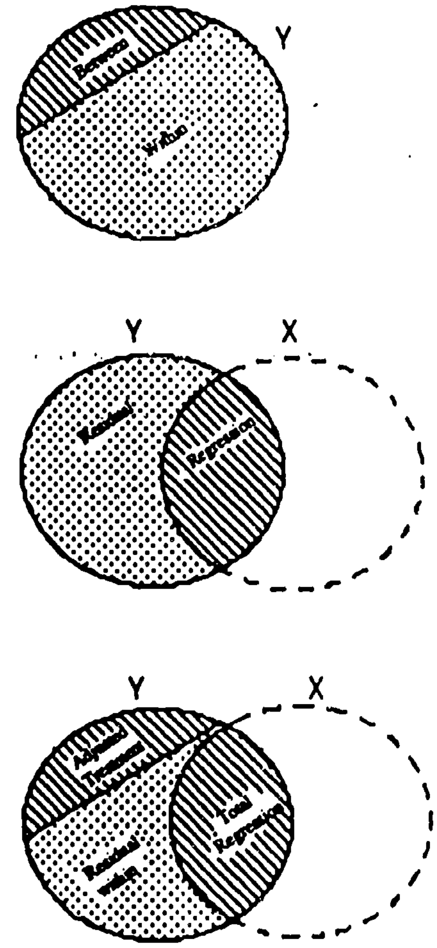

Figure 2 Comparison of ANOVA, ANOVAR, and ANCOVA when ANCOVA assumptions are met. (Adapted from Huitema, 1980, p. 26)

The linear model for the ANCOVA, $Y_{ij} = \mu + \alpha_j + \beta_1 (X_{ij} - \bar{X}..) + e_{ij}$, illustrates that this procedure is actually an ANOVA on *adjusted scores* (Hinkle, Wiersma, & Jurs, 1988) where

Y_{ij} = dependent variable score of *i*th individual in *j*th group
μ = population mean (on dependent variable) common to all observations
α_j = effect of treatment *j* (a constant associated with all individuals in treatment *j*)
β_1 = linear regression coefficient of Y on X
X_{ij} = covariate score for *i*th individual in *j*th group
$\bar{X}..$ = mean of all individuals on covariate
e_{ij} = error component associated with *i*th individual in *j*th group

The $\beta_1(X_{ij}-\bar{X}..)$ is the adjustment for each individual score. The ANCOVA variance partition shows the treatment effects that are independent of X (the adjusted treatment), the differences in achievement among subjects that can be predicted from test X, and the differences among subjects that are not due to treatment effects and cannot be predicted from test X (i.e., error) (Huitema, 1980). These three parts of the variance make up the sums of squares total. The sum of squares (SS) total, the SS between and the SS within are adjusted for the variance attributed to the covariate. The total regression area represents variability on Y predictable from X. If that area is removed, then the remaining area contains only variability due to treatment effects independent of X and variability due to error (Huitema, 1980). This remaining area is referred to as the SS residual total.

The SS within groups represents differences predictable from X and differences not predictable from X. By subtracting the SS due to predictable differences among subjects within groups from SS within groups, the SS residual is obtained to then be used as the SS error in ANCOVA. The SS error accounts for differences that are not predictable from X and are not accounted for by treatment differences. In Figure 1, the SS residual within is the SS error.

The SS adjusted treatments effects (see Figure 2, ANCOVA) is obtained by subtracting the SS residual within from the total residual SS (Huitema, 1980). This area is the quantity that was referred to as the treatment effects independent of X. The ratio obtained from the SS adjusted treatment and the SS residual total is the " descriptive measure of the proportion of the variability explained by the treatments when the effects of the covariate are controlled statistically" (Huitema, p. 31).

ANCOVA sounds like it would offer a great statistical solution to the researcher's problems. As Loftin and Madison (1991, p. 134) point out, "the 'correction' of the dependent variable scores is seen by some as a device to adjust for all kinds of problems, but very few data sets can meet the very specific requirements that make the adjustment appropriate"and Thompson (1991) states that "the inappropriate use of ANCOVA appears to be the rule rather than the exception".

What is "inappropriate use"? Keppel and Zedeck (1989) repeatedly state that ANCOVA is only appropriate for use with **randomly assigned groups** and then only after the **homogeneity of regression assumption** has been met. These authors note:

> It is somewhat depressing to note that while all statistical methodology books continue to stress [would that this were true] the conclusion that ANCOVA should not be used in quasi-experimental designs, misapplications of the procedure are still committed and reported in the literature. (p. 482)

Campbell and Erlebacher (1975) cite several authors who have warned of the dangers of ANCOVA, but other authors give wrong recommendations on correcting the serious problems with ANCOVA. Campbell and Erlebacher discuss the article by Williams and Evans (1969) describing the analysis of the Head Start program: "The one weakness in Williams and Evans' otherwise outstanding paper comes at this point. They state that ex post facto studies are a respected and widely used scientific procedure" (p.613). Campbell and Erlebacher say that in actuality:

> most methodology texts are silent on the issue and condone comparable procedures...On using analysis of covariance to correct for pretreatment differences, the texts that treat the issue are either wrong or noncommittal (that is, they fail to specify the direction of the bias), and probably 99% of the experts that know of the procedure would make the error of recommending it. (p. 613)

In the Head Start program analysis, the analysis of covariance made an underadjustment and made the Head Start program seem as if it actually damaged low socioeconomic children (Campbell&Erlebacher, 1975). Thompson (1989) notes that because of the inappropriate use of ANCOVA the Head Start program was almost discontinued.

What are the conditions required for correct ANCOVA usage? , Loftin and Madison (1991, p. 134) state the following conditions that must be met (as referred to at the beginning of this paper):

1. The covariate (or covariates) should be an independent variable highly correlated with the dependent variable.
2. The covariate should be uncorrelated with the independent variable or variables.
3. With respect to the dependent variable, (a) the residualized dependent variable (Y*) is assumed to be normally distributed for each level of the independent variable, and (b) the variances of the residualized dependent variable (Y*) for each level of the independent variable are assumed to be equal.
4. The covariate and the dependent variable must have a linear relationship, at least in conventional ANCOVA analyses.

5. The regression slopes between the covariate and the dependent variable must be parallel for each independent variable group.

This paper is concerned with assumption 5 and will describe and illustrate the assumption of the homogeneity of regression as it relates to ANCOVA with three treatment groups.

Homogeneity of Regression

Homogeneity of regression means that the slopes for each of the three groups, when the regression equations are computed individually within groups, must be parallel. Figure 3 below shows an actual regression plot of the data in Table 1 with an individual line for each group. In order for the homogeneity of regression assumption to be met, the relationship between X and Y has to be the same for each of the three groups. Figure 3 shows that the slopes are roughly the same, therefore we could be confident that any adjustment made using the covariate would result in the same adjustment for each group.

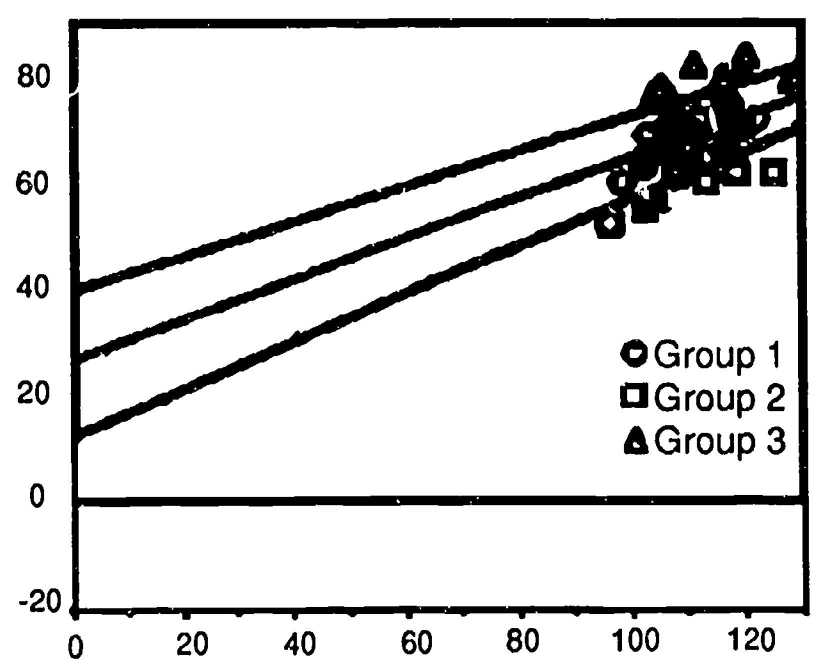

Figure 3 Illustration of a data set that meets the homogeneity of regression assumption

Why is it so important that the three groups have the same regression slope? Loftin and Madison (1991) state:

> If the regression slopes are equal, a single pooled regression slope may be used for all the subjects , regardless of their group assignment, to calculate the solicited adjustments in the dependent variable (Y). And ANCOVA always uses a pooled equation created under an assumption that since the equations are supposed to be the same in the different groups, an "average" equation can be used for all the subjects regardless of their groups. (p. 141)

The ANCOVA uses a pooled regression line when making the adjustments and totally ignores the regression slope for each group. To test the homogeneity of regression mathematically a residual sums of squares is computed for each group and then compared to the residual sums of squares using a pooled regression equation that ignores group membership. If these are equal, the slopes are said to be homogeneous and "the adjusted means are descriptive measures because the treatment effects are the same at different levels of the covariate" (Huitema, 1980, p. 43).

An F test can be done to see if the difference in the slopes is statistically significant or not. This can be helpful when a scatterplot of the data results in a graphic representation such as Figure 4. Does this data set meet the homogeneity of regression assumption or not? Using an F test, yes, these data (see Table 2) do meet the homogeneity of regression assumption. Of course, we should keep in mind though that an F test is influenced by sample size and that we would get statistical significance if we had a large enough sample size.

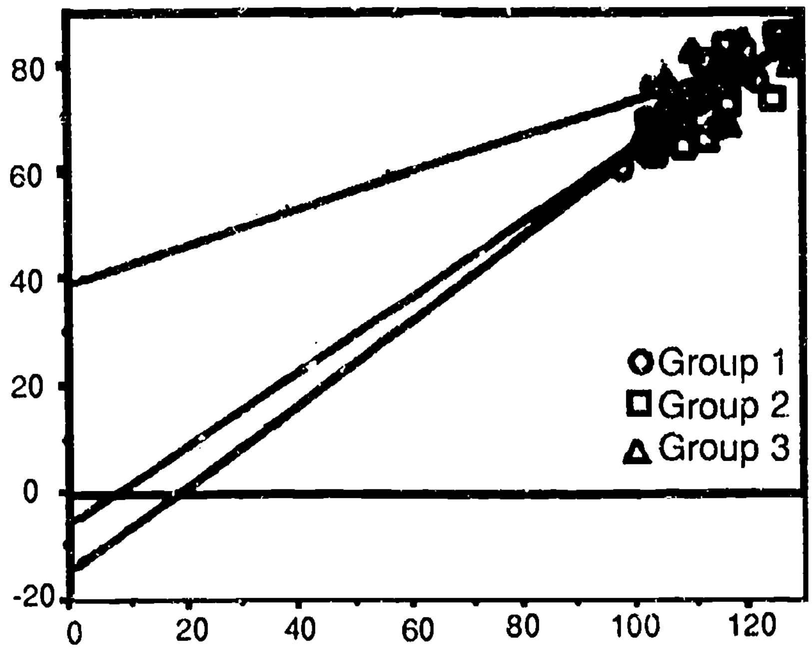

Figure 4 Illustration of a data set that meets the homogeneity of regression assumption using an F test

What steps are actually involved in doing the F test? After computing the within-group residual sum of squares (SS_{res_w}), individual sums of squares residual are computed for each group and added together to obtain a sum of squares residual individual (SS_{res_i}). The next step is to compute the heterogeneity of slopes sum of squares. This is done by subtracting the (SS_{res_i}) from the (SS_{res_w}). This difference "reflects the extent to which the individual regression slopes are different from the pooled within-group slope" (Huitema, p. 45). Huitema (1980) also points out:

> that SS_{res_i} can never be larger than SS_{res_w} , just as, in an ordinary ANOVA, the sum of squares within can never be larger than the sum of squares total. There is only one explanation for SS_{res_w} being larger than SS_{res_i} - the individual within-group slopes must be different. (p. 45)

Figure 5 illustrates the heterogeneous slope case which was generated from the data set in Table 3.

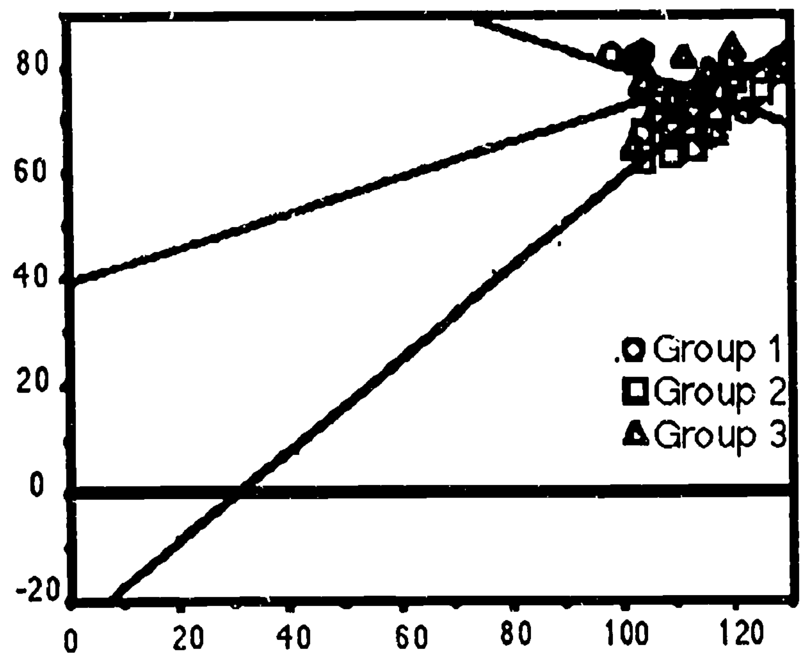

Figure 5 Illustration of a data set that has heterogeneous slopes

When the individual slopes of the lines are the same, they are also equal to the pooled regression line slope, but when the slopes differ, as in Figure 5, the pooled regression slope line cannot possibly have residuals as low as the individual regression lines. Stated differently, if the individual slopes are equal, the points fall around each of the slope lines in the same way and the pooled regression line represents an average of the residuals. But if the points do **not** fall around the slope lines in the same way (creating different slopes) then an average of the three slopes cannot possibly be an accurate description or representation of the three different groups. If we were to draw the pooled regression line on Figure 5, we would see that the points would be much further away from this pooled line than they are from the individual regression lines. Huitema (1980, p. 46) states, "A single regression slope simply cannot fit different samples of data as well as can a separate slope for each sample, unless there are no differences among the slopes."

What happens during the ANCOVA correction that makes the homogeneity of regression assumption important? All of the Y scores are adjusted into new Y scores (Y*) by the following formula:

$$\bar{Y}^*_k = \bar{Y}_k - b_w (\bar{X}_k - \bar{X})$$

where

$\bar{Y}^*_k$ = adjusted group mean on the dependent variable

$\bar{Y}_k$ = unadjusted group mean on the dependent variable

$\bar{X}_k$ = group mean on the covariate

$\bar{\bar{X}}_k$ = grand mean on the covariate

b_w = pooled within-groups regression coefficient reflecting the correlation between the dependent variable and the covariate

The pooled regression coefficient is used in the correction and a straight line relationship between X and Y is assumed. Loftin and Madison (1991) emphatically state:

> Since a single regression equation is used to correct all Y scores, regardless of independent variable groups, if the "pooled or "averaged" single equation does not accurately describe the Y and X relationship in a given group, the corrections producing the residual Y scores...will actually bias the data rather than "correct" them. (p. 143)

Summary

One consequence of violating the homogeneity of regression assumption is biased F tests. If the individual regression lines for each group are not parallel, then the pooled regression line does not fit and is not a good average for the three groups. As the heterogeneity of the slopes increases, the pooled regression line is less able to represent the individual groups. So the error sum of squares increases. Furthermore, when the homogeneity of regression assumption is not met the residualization using the covariate may actually reduce the sum of squares for the intervention, thus artificially making it look less effective or even completely ineffective. Consequently, in the F test, the more heterogeneous the slopes are the smaller the ANCOVA F will be. Since smaller F values are associated with larger probabilities, then the bias occurs because the null hypothesis is rejected when it should not have been.

An ANCOVA has its place in limited situations and **only** after meeting necessary assumptions. For these very reasons other methods of analysis should be considered seriously before an analysis of covariance when intact groups are employed in a study.

References

Campbell, D.T., & Erlebacher, A. (1975). How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In M. Guttentag & E.L. Struening (Eds.), Handbook of evaluation research (Vol. 1, pp. 597-617). Newbury Park: SAGE.

Keppel, G., & Zedeck, S. (1989). Data analysis for research designs. New York: W.H. Freeman.

Hinkle, D. E., Wiersma, W. & Jurs, S.G. (1988) Applied statistics for the behavioral sciences. Boston: Houghton Mifflin Company.

Huitema, B. E. (1980). The analysis of covariance and alternatives. New York: John Wiley & Sons.

Loftin, L.B. & Madison, S. Q. (1991). The extreme dangers of covariance corrections. In B. Thompson (Ed.), Advances in educational research: Substantive findings, methodological developments (Vol. 1, pp.133-147). Greenwich, CT: JAI Press Inc.

Thompson, B. (1989). The place of qualitative methods in contemporary social science. In B. Thompson (Ed.), Advances in social science methodology (Vol. 1, pp. 1-49). Greenwich, CT: JAI Press.

Thompson, B. (1992). Two and one-half decades of leadership in measurement and evaluation. Journal of Counseling and Development, 70, 434-438.

Table 1

Group 1 - IQ	Group 2 - IQ	Group 3 - IQ	Group 1 - Ach.	Group 2 - Ach	Group 3 - Ach
98	96	102	60	55	65
102	102	117	63	58	68
104	104	108	66	59	72
103	110	117	69	67	76
112	108	105	72	64	78
113	113	116	75	63	80
118	117	111	71	75	82
120	110	120	70	70	84
115	113	107	67	68	75
106	109	104	70	69	77
112	125	128	69	65	79
122	118	119	72	65	81

Table 2

Group 1 - IQ	Group 2 - IQ	Group 3 - IQ	Group 1 - Ach.	Group 2 - Ach	Group 3 - Ach
98	104	102	60	62	65
102	109	117	63	63	68
104	104	108	66	67	72
103	117	117	69	71	76
112	120	105.	72	77	78
113	113	116	75	79	80
118	117	111	78	82	82
120	126	120	80	84	84
115	113	107	67	64	75
106	109	104	70	68	77
112	125	128	74	72	79
122	118	119	76	77	81

Table 3

Group 1 - IQ	Group 2 - IQ	Group 3 - IQ	Group 1 - Ach.	Group 2 - Ach	Group 3 - Ach
98	104	102	[illegible]	62	65
102	109	117	81	63	68
104	104	108	83	67	72
103	117	117	82	71	76
112	120	105	72	77	78
113	113	116	75	71	80
118	117	111	78	70	82
120	126	120	80	88	84
115	113	107	67	64	75
106	109	104	70	68	77
112	125	128	74	75	79
122	118	119	72	74	81