MULTIPLE LINEAR REGRESSION VIEWPOINTS
VOLUME 17, NUMBER 1, FALL 1989

The Meta-Analysis of the Effect of Class Size on Achievement: A Secondary Analysis

Susan M. Tracz
California State University at Fresno

Dennis W. Leitner
Southern Illinois University at Carbondale

Abstract

One of the first examples of the use of Gene Glass' meta-analysis was the Glass and Smith studies of the effect of class size on achievement in school. It was concluded that "a clear and strong relationship between class size and achievement has emerged" (Glass & Smith, 1979). This paper presents the reanalysis of the Glass and Smith data, removing small classes of five or less, which are virtually tutorial sessions. The results show a greatly reduced effect on achievement for small classes.

An earlier version of this paper was presented at the annual meeting of the Mid-Western Educational Research Association in Chicago, Sept. 27-29, 1984.

Introduction

If teachers were asked if they favored smaller classes over larger ones, the vast majority would probably respond that they favored smaller classes (Bain & Achilles, 1986). The rationale for this might be expressed in the following ways: the teacher-student rapport is better in smaller classes, teachers can individualize instruction to a greater extent resulting in greater learning in smaller classes, and the attitudes of both teachers and students improve in smaller classes. The importance of small classes can be underscored by noting that this topic is often an issue in teacher contract negotiations. The opposing position, usually held by school administrators, is that the achievement of students in larger classes is equivalent to that of students in smaller classes and the larger classes are more cost-effective.

Although a considerable number of research studies have compared student achievement in small versus large classes, a representative sampling of the literature would lead to inconclusive findings: studies can be found that favor large classes and other studies can be found that indicated an advantage to small classes. Therefore, this topic is an ideal one for the application of a statistical technique called meta-analysis.

Meta-analysis, pioneered by Gene V Glass, is a statistical methodology for integrating a large number of individual studies. Glass (1976) divided research into two types: primary analysis and secondary analysis. He defined primary analysis as "original analysis of data in a research study," while secondary analysis is defined as "re-analysis of the data for the purpose of answering the original research question with better statistical techniques or answering new questions with old data" (p. 3). He continued to propose a new type of analysis, meta-analysis, which "refers to the analysis of analyses...[or] the statistical analysis of a large collection of analyses results from individual studies for the purpose of integrating the findings" (p. 3).

The results of a meta-analysis are often presented in terms of mean effect size and its place on the normal distribution. Effect size is usually defined either as the difference between means of experimental and control groups divided by a standard deviation:

$$ES = \frac{(\bar{X}_E - \bar{X}_C)}{s}$$

where s = either the standard deviation of the control group
or a pooled estimate of the standard deviation

or as a correlation coefficient:

$$ES = r.$$

Glass and Smith's Original Meta-analysis

Glass and Smith (1976) performed a meta-analysis on the relationship between class size and achievement. Their estimate of effect size was given by:

$$ES_{SL} = \frac{(\bar{X}_S - \bar{X}_L)}{s}$$

where $\bar{X}_S$ = the mean achievement for the smaller class,
$\bar{X}_L$ = the mean achievement for the larger class, and
s = the estimated pooled, within-class standard deviation

After a careful search of the previous studies of the class size literature, the document retrieval and abstracting resources, and the bibliographies of the studies which were found, 77 studies were identified which yielded 725 effect sizes. Glass and Smith (1979) reported that the mean of the 725 effect sizes was .088 and the median was .050. The standard deviation was .401, the skewness 1.151 and the kurtosis 7.461. The effect sizes ranged from -1.98 to 2.54, and 40% were negative while 60% were positive (i.e. favoring smaller classes).

Glass and Smith (1979) fit the following quadratic least squares regression model to the data:

$$ES_{S-L} = \beta_0 + \beta_1 S + \beta_2 S^2 + \beta_3 (L - S) + \varepsilon$$

where S = the size of the smaller class,
L = the size of the larger class,
$\beta_0, \beta_1, \beta_2, \beta_3$ = the population regression weights, and
ε = the error of estimate

Glass and Smith (1979) obtained the following summary table:

Source of Variation	df	MS	F
Regression	3	6.684	50.636
Residual	721	.132	

The multiple R for the model was .426. Substituting the estimated regression weights in the model yielded the following regression equation:

$$\hat{ES}_{S\text{-}L} = .57072 - .03860\,S + .00059\,S^2 + .00082\,(L - S)$$

A graph of the regression line for achievement in percentile ranks on class size for all data appears in Figure 1 (Note from Glass & Smith, 1979).

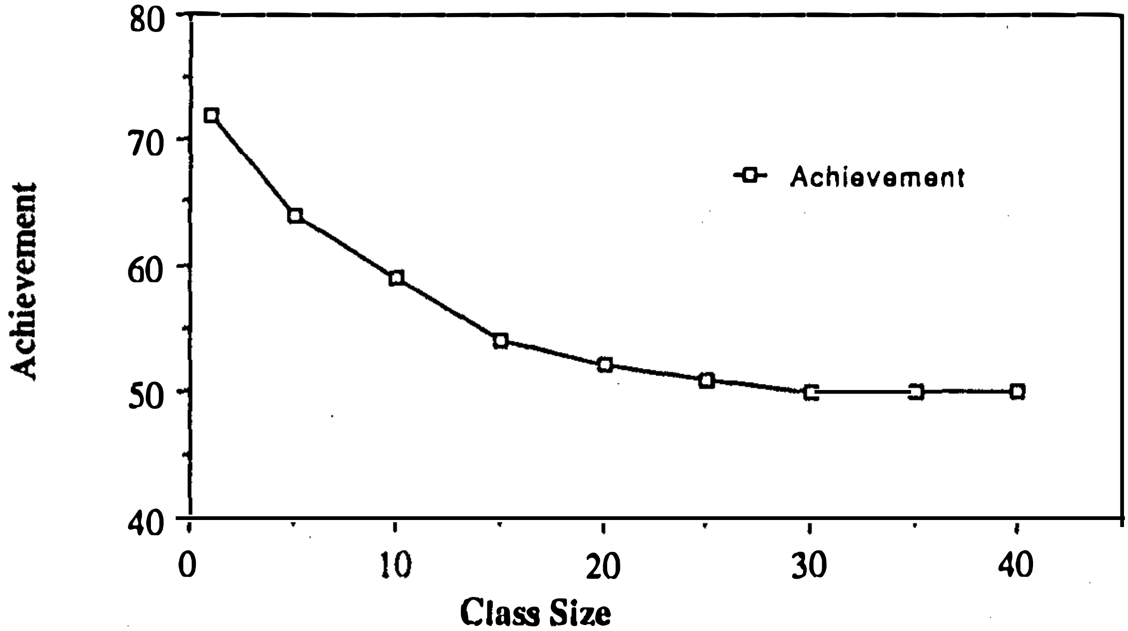

Figure 1. Regression Line for Achievement

A number of other variables were also analyzed in this meta-analysis. Included among these were year of the study, duration of instruction, pupil/instructor ratio, pupil ability, age, assignment of pupils and teachers, type of achievement measure and quantification of outcomes. However, of all the regression analyses performed on the data, only two analyses provided any meaningful information. These analyses were based on two comparisons: elementary vs. secondary students and well-controlled vs. poorly-controlled studies.

Students were sorted by age into two groups: those who were 11 years old or younger (elementary students) and those who were 12 years old or older (secondary). Separate regression analyses using the model given earlier yielded the following results for elementary school-aged children:

ELEMENTARY (N=342)

Source of Variation	df	MS	F
Regression	3	1.898	38.735
Residual	338	.049	

The multiple R for this model was .505. Substituting the estimated regression-weights into the model yielded the following equation:

$$\hat{ES}_{S\text{-}L} = .38503 - .02995\,S + .00052\,S^2 + .00344\,(L - S)$$

The following results were obtained for secondary school-aged pupils:

SECONDARY (N=349)

Source of Variation	df	MS	F
Regression	3	5.667	27.377
Residual	345	.207	

The multiple R for this model was .439 and the regression equation was given by the following:

$$\hat{ES}_{S\text{-}L} = .75539 - .05024\,S + .00071\,S^2 + .00111\,(L - S)$$

A graph of the regression lines for both the elementary and secondary groups for achievement in percentile rank on class size appears in Figure 2 (Note from Glass & Smith, 1979). The graph indicates that the relationship between small class size and higher achievement is more pronounced in the secondary grades than in the elementary grades.

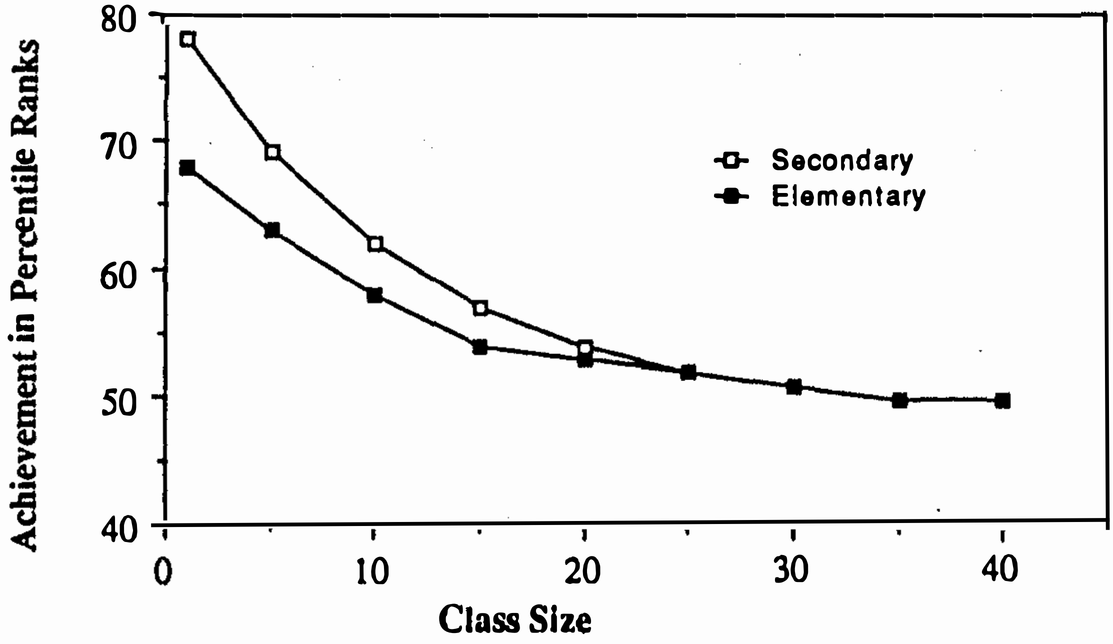

Figure 2. Regression of Achievement onto Class Size by Grade Level

Finally, comparable regression analyses were done on groups of studies classified as well-controlled versus studies classified as poorly controlled. In well-controlled studies, students were randomly assigned to large and small classes, while intact classes were used in the poorly-controlled studies.

The analysis of well-controlled studies provided the following results:

WELL-CONTROLLED (N=108)

Source of Variation	df	MS	F
Regression	3	4.226	21.784
Residual	104	.194	

The multiple R for this model was .621. Substituting the estimated regression weights into the model yielded the following equation:

$$\hat{ES}_{S\text{-}L} = .69488 - .06334\,S + .00128\,S^2 + .00783\,(L - S)$$

The analysis of the poorly-controlled studies yielded the following results:

POORLY-CONTROLLED STUDIES (N=334)

Source of Variation	df	MS	F
Regression	3	.263	3.985
Residual	330	.066	

The multiple R for this model was .187 and the regression equation was given by:

$$\hat{ES}_{S\text{-}L} = .07399 - .00587\,S + .00009\,S^2 + .00376\,(L - S)$$

A graph of the regression lines for both the well-controlled and poorly-controlled studies for achievement in percentile ranks appear in Figure 3 (Note from Glass & Smith, 1979). For studies using random assignment of student, the achievement in small classes was markedly higher than in the poorly controlled studies where random assignment was not used.

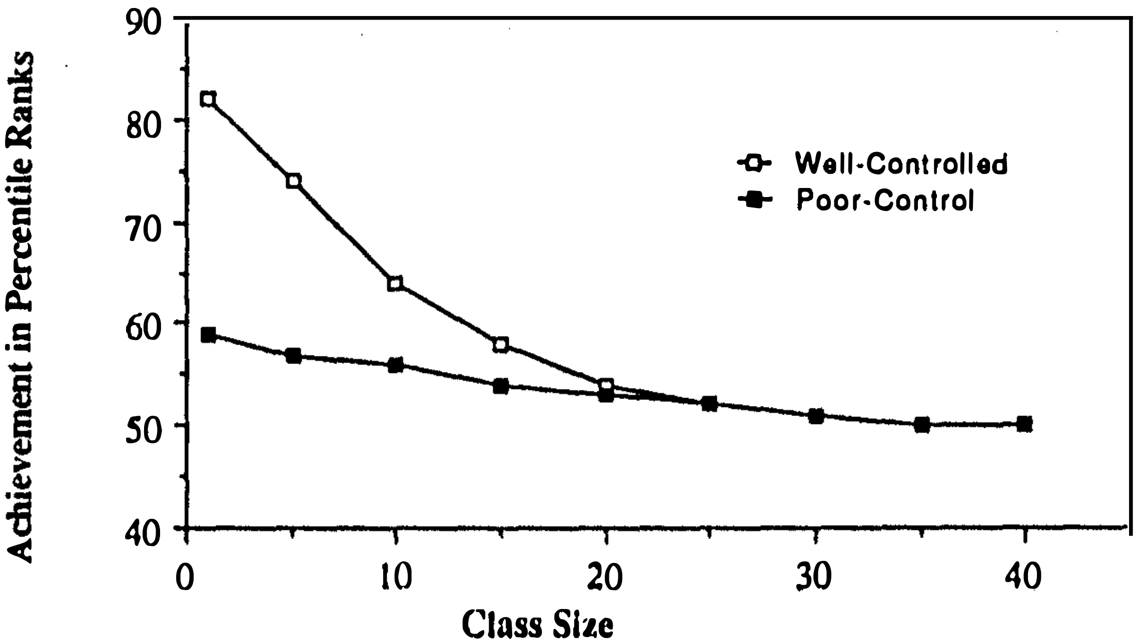

Figure 3. Regression of Achievement onto Class Size by Control

Glass and Smith (1979) concluded that:

> "a clear and strong relationship between class size and achievement has emerged. The relationship seems slightly stronger at the secondary grades than the elementary grades, but it does not differ appreciably across different schools subjects, level of pupil IQ, or several other obvious demographic features on classrooms. The relationship is seen most clearly in well-controlled studies in which pupils were randomly assigned to classes of different sizes" (p. 15).

Criticisms of the Glass and Smith Meta-analysis

Although the work of Glass and Smith appears to be quite conclusive, it has not gone without criticism. First, one contradiction in their findings is that only 60% of the effect sizes were positive although they claim a "clear and strong relationship between class size and achievement" exists. This means that, in nearly half (40%) of the effect sizes, the achievement of the larger class exceeded the achievement of the smaller class. In addition, an R^2 of .181 leaves almost 82% of the variance of achievement unexplained by variation in class size. Even though a highly significant proportion of variance is accounted for, there is much room for improvement.

Another criticism, presented by the Educational Research Service (1980), was that the graph of achievement regressed on class size for well-controlled studies vs. poorly-controlled studies was based on only 14 studies. Of these 14 studies, a mere six studies were conducted in situations that are typical of elementary and second school.

Perhaps the most telling criticism of all pertains to the range of class sizes where the largest increments in achievement occur. As all the graphs presented illustrate, the most pronounced change in the rate of achievement occurs in classes smaller than 15 in number. Only minimal differences in achievement can be seen in the range of 20 to 40 students, which are the more typical sizes of classes.

Reanalysis Eliminating Tutorials

A large number of the small classes had only one to five students enrolled. These classes could more accurately be called "tutorial sessions." The purpose of this study was to reanalyze the Glass and Smith data eliminating the very small, atypical, class sizes and observing the resulting effect sizes to see the impact of the "tutorial sessions." The data reported by Glass and Smith (1978) were entered into the computer for this reanalysis. The results appear in Table 1.

Table 1. Means and Standard Deviations of models of Effect Sizes for Original and Successively Reanalyzed Data from Glass and Smith (1978)

Study	N	Mean	St. Dev.	R^2	p <
Glass and Smith	725	.088	.401	.181	.0001
Tracz and Leitner	662	.091	.406	.180	.0001
Eliminating effect sizes based on small class of					
1	609	.046	.356	.060	.0001
1-2	607	.045	.356	.058	.0001
1-3	601	.033	.332	.017	.0176
1-4	599	.031	.330	.013	.0493
1-5	598	.031	.330	.012	.0609

Although Glass and Smith (1978) appear to have presented their entire data set, the listing omits 63 effect sizes. For the available data, the mean was .091. The standard deviation of the two data sets was almost identical as was the multiple R^2 for our reanalysis. This gave us confidence, that while some studies were missing, our reanalysis was not substantively affected.

When the 53 effect sizes that included the small classes with only one student were removed from the analysis, the mean effect size dropped from .091 to .046 - a decrease of nearly 50%. The standard deviation dropped from .406 to .356 and the R^2 dropped from .180 to .060 - leaving 94% of the variance unaccounted for! The 53 effect sizes were from 8 studies, averaged .608 with a standard deviaiton of .566 and ranged from -.44 to 2.52.

When an additional nine effect sizes were eliminated, representing small classes with two through five students, the mean decreased even further to .031, approximately one-third of what it was for the full data set. The model involving the three variables used in all analyses accounts for slightly more than 1% of the variance in Effect Size and is no longer significant at the .05 level.

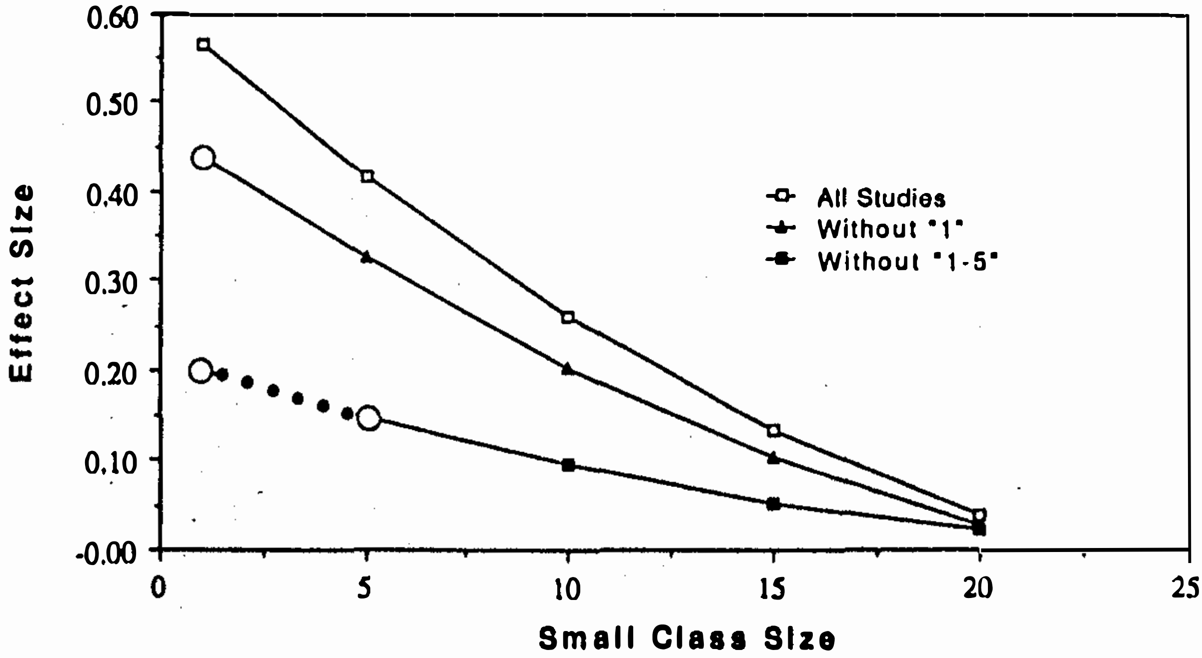

Figure 4. Effect Sizes With and Without Small Classes

Figure 4 is a graph of the three regression of effect size on class size, plotted for the small class sizes of 1 to 20 (assuming a large class size of 38, which is the average large class size for all studies). The general regression equation is:

$$ES_{S\text{-}L} = \beta_0 + \beta_1 S + \beta_2 S^2 + \beta_3 (L - S) + \varepsilon$$

Looking at the left side of the graph, the top line represents the regression line for all the data, the middle line represents the regression line with the small classes of one student eliminated from the analysis, and the bottom line represents the regression with small classes of fewer than six pupils eliminated from the analysis. Empty circles and dotted lines depict projected information where data were eliminated (i.e., studies with small class sizes of 1-5 removed). When the unrealistically small classes are removed, the predicted effect size dramatically decreases. The predicted effect size for a class of one student drops from approximately .58 to .21, from what one, using Cohen (1977), might call a drop from a medium to a small effect size. When the small class consists of about 20 students, the effect size is about .05. If the average of a class of 38 were considered to fall at the 50 percentile, a class of 20 would fall at the 52nd percentile (which is the percentile rank of a z-score of .05).

The fact that class sizes for five or fewer students are virtually impossible for the vast majority of school districts is underscored by the tenor of the major longitudinal research study conducted in and partially funded by the state of Tennessee. The researchers conducting this study, Whittington, Bain and Achilles (1985) state that

> ...class size studies have often investigated the wrong sizes, studying reductions from 36 to 25 pupils are various grade levels. Perhaps the real payoffs are achieved by reducing class size significantly — to 15 pupils per classroom teach in the primary grades. (p. 33)

This rigorously conducted, three-year study followed students from kindergarten, compared achievement in classes of 15 and 25 students and found significantly higher achievements in the smaller classes (Bain, Achilles, & Witherspoon-Parks, 1988). However, those smaller classes were much larger than many included in Glass and Smith's meta-analysis, and it is the effect sizes from these extremely small classes that drastically inflate the mean effect size they report.

In conclusion, the increased achievement that Glass and Smith attributed to small classes may be substantially less than claimed after deleting the effect sizes based on atypically small classes of one to five students. However, other positive by-products of small but feasible class sizes may still be found.

References

Bain,. H. P.& Achilles, C. M. (1986). Interesting developments on class size. *Phi Delta Kappan, 67*(9), 662-665.

Bain, H. P., Achilles, C. M., & Witherspoon-Parks, M. (1988, November). *Three-year Longitudinal study of small class size: The Metro-Nashville public schools study: 1984-1987.* Paper presented at the Mid-South Educational Research Association, Louisville, KY.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences, (Rev. ed.)* New York: Academic Press.

Educational Research Service. (1980). Class size research: Critique of recent meta-analysis. *Phi Delta Kappan, 62*, 239-241.

Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational Researcher, 5*, 3-8.

Glass, G. V & Smith, M. L. (1978). *Meta-analysis of research on the relation ship of class-size and achievement.* San Francisco: Far West Laboratory for Educational Research and Development.

Glass, G. V & Smith, M. L. (1979). Meta-analysis of research on class size and achievement. *Educational Evaluation and Policy Analysis, 1*, 2-16.

Whittington, E. H., Bain, H. P. & Achilles, C. M. (1985). Effects of class size on first-grade students. *Spectrum: Journal of School Research and Information, 3*(4), 33-39.