ERS Special Report

CLASS SIZE RESEARCH:
A Critique of Recent Meta-Analyses

Educational Research Service, Inc.

ERS Special Report

CLASS SIZE RESEARCH:
A Critique of Recent Meta-Analyses

Educational Research Service, Inc.

Educational Research Service, Inc.
1800 North Kent Street
Arlington, Virginia 22209
Phone: (703) 243-2100

The Information Source for School Management

Sponsored by

- American Association of School Administrators
- American Association of School Personnel Administrators
- Council of Chief State School Officers
- National Association of Elementary School Principals
- National Association of Secondary School Principals
- National School Boards Association
- National School Public Relations Association

Educational Research Service is an independent, nonprofit corporation serving the research and information needs of the nation's school systems, their administrative teams, school boards, and the public.

ERS serves as both a national source and clearinghouse for school management research and information. By compiling, analyzing, and sharing information that is essential to effective decision making, ERS assists local school districts in both day-to-day operations and long-range planning. It provides an invaluable approach to data gathering and information reporting through services that are geared to quick yet thorough response to requests from subscribers. An ERS subscription entitles the subscribing school system or agency to: • on-call information service for its administrative staff and governing board • an initial copy of each ERS conducted research study produced during the subscription year • the monthly *ERS Bulletin* summarizing the research of other agencies.

To assure that small as well as large systems can benefit from ERS, a graduated subscription fee has been established. This enables subscribing agencies to share equitably in the cost of compiling and reporting the research needed by all school administrators and governing boards. ERS services are available at reasonable rates to associations of school administrators and school boards, as well as university departments of school administration. Services may be provided to other groups by special action of the ERS Board of Directors. Subscription rates are available upon request.

ERS Executive Staff

Glen E. Robinson
President
Director of Research

John M. Forsyth
Vice President

Price of publication: $12.00. To ERS subscribers: single copy mailed routinely as part of the subscription; additional copies, half price.
Stock #224-00002

CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

This critique of two recent meta-analyses of class size research is unique among all previous studies and reports published by the Educational Research Service. It has been made necessary by two extraordinary reviews of class size research prepared by Gene V Glass and Mary Lee Smith and published by the Far West Laboratory for Educational Research and Development under a federal grant. These two reports have unusual importance since they hold the possibility of having far-reaching yet unwarranted impact on a whole array of educational policy and management decisions pertaining to the assignment of pupils, teachers, and resources to classrooms.

In their first report, Glass and Smith announced that their findings contradicted the conclusions reached by previous analysts of class size research. Because of their use of "sophisticated" meta-analysis (a process which Glass described as an "analysis of analyses") and the large volume of class size research studies examined, the authors claimed that their findings were superior to all previous reviews of the class size research. The authors stated that their techniques enabled them to make "bold generalizations" about the effects of class size on pupil achievement where other analysts could make only "timid qualifications." Based primarily on a graph developed from the findings of 14 class size studies which the authors said exercised good experimental control, Glass and Smith declared unequivocally that their analysis "established clearly

that reduced class-size can be expected to produce increased academic achievement." [16:iv]*

Since the meta-analyses were first released, a number of individuals and organizations across the country have interpreted these findings as conclusive proof that smaller classes, without question, result in superior effects on both achievement and nonachievement aspects of pupil instruction. In an initial report of the first meta-analysis, the journal *Phi Delta Kappan* stated: "The Glass study is the first by a nationally recognized researcher to make unequivocal statements about the effects of class size on pupil achievement. *It has enormous policy implications.*" [emphasis added] [18:411]

In view of such unprecedented conclusions and possible impact, ERS considered it important to make a thorough examination of the two meta-analyses. In the course of this examination, ERS was surprised and concerned to find that several of the claims and conclusions made in the two Glass and Smith meta-analysis reports were unsupported. Moreover, a number of the interpretations of the meta-analysis findings and the recommendations for educational policy based on these conclusions were not only unsupported but also conflicting. Thus, as the examination continued, it became clear that in light of the magnitude of harm that could occur, ERS was obligated to publish a full critique of the two meta-analyses.

The purpose of this critique, therefore, is to report to persons and agencies concerned with class size issues the results of the examination made by ERS of the methodology, findings, and conclusions of the two Glass and Smith meta-analyses and published interpretations of them. It should be *emphasized* from the beginning that the purpose of this critique is *not* to make the case in favor of either smaller classes or larger classes. Rather, the intent is to provide a reliable analysis of the information contained in the two recent meta-analyses. In doing this, it is hoped that school administrators, school board members, state legislators, teachers, parents, and other interested parties may be better able to determine when, where, and with what groups of pupils and teachers smaller classes could be beneficial.

**References cited in the body of the text are noted by numbers within brackets. The number before the colon indicates the entry number within the list of references on page 79; the number following the colon indicates the page within the entry. Where no colon appears, the citation refers to the entire entry. Multiple citations are separated by semicolons.*

Because of the many and far-reaching instructional, financial, and political implications of the two meta-analyses, it is necessary that this critique examine fully the methods, findings, and implications of these two class size reports. Such treatment makes the critique longer and somewhat more technical than would be otherwise desirable, but in view of the "enormous" possible implications of the two meta-analyses, full treatment is essential if this critique is to provide school officials and others with important information they should have when making class size decisions.

Background

The research and related literature on the subject of class size is immense. It covers the effect of class size on pupil achievement, pupil attitudes, teacher morale and attitudes, classroom methods, policy decisions, and school district finances. In the past 25 years, at least 23 reviews of the research on class size have been published; three-fourths of these have appeared in the 1970s alone. (See Appendix C beginning on page 75.)

Reviewers of this research generally have concluded that the relationship between class size and pupil achievement is inconclusive--with some studies finding that smaller classes are better, some that larger classes are better, and some reporting that there is no difference between the two. There is no argument that measuring class size effects is a complex task and that numerous other variables interact with class size. It is also widely recognized that in general teachers perceive an inherent value in smaller classes, both in terms of pupil effects and teacher morale, and that class size policy decisions can have far-reaching political and financial repercussions.

Educational Research Service examined the literature on class size and in May 1978 published its findings in an ERS Research Brief titled *Class Size: A Summary of Research*. [6] Among the topics discussed in this analysis were the effects of class size on pupil achievement in the elementary and secondary grades, the effects of class size on the classroom environment, teacher opinions regarding class size, the concept of "optimum" class size, and some of the policy implications related to the class size issue. This Research Brief also offered 19 *tentative* conclusions for consideration by school policy makers; these conclusions are reprinted in their entirety in Appendix B beginning on page 71.)

The ERS Research Brief acknowledged that policy decisions relating to class size involve factors that are complex, varied, and often emotionally charged. However, in terms of pupil benefits, ERS found that class size research does not support the contention that smaller classes *alone* will produce increased academic achievement. Within the mid-range of about 25 to 34 pupils, ERS concluded, class size appears to have little effect on pupil achievement for most pupils in most subjects above the primary grades. ERS found that smaller classes *can* have a positive impact on pupil achievement in reading and mathematics in the early primary grades and for pupils with lower academic ability and for economically or socially disadvantaged pupils. [6:69]

After the release of the ERS Research Brief on class size, the Far West Laboratory for Educational Research and Development, located in San Francisco, published two summaries of the research on class size. The first, issued in September 1978, was titled *Meta-Analysis of Research on the Relationship of Class-Size and Achievement.*[1/] [16] The authors of this study were Gene V Glass and Mary Lee Smith, both of the Laboratory of Educational Research at the University of Colorado. The second class size study, titled *Relationship of Class-Size to Classroom Processes, Teacher Satisfaction and Pupil Affect: A Meta-Analysis,*[2/] was released in July 1979. [30] This study also was authored by Smith and Glass.[3/]

These two studies were the first reports published in conjunction with the Far West Laboratory's Class Size and Instruction Project. In addition to the two meta-analyses, the Project has conducted field studies in two second grade classrooms in a rural Virginia school and two in an inner-city California school to investigate the effects of actual class size reduction in the classroom. A description of this study was published by the Far West Laboratory in a January 1980 report titled *What Happens in Smaller Classes? A Summary Report of a Field Study.* [10]

[1/] This study will be referred throughout the critique as "Meta-Analysis I."

[2/] This study will be called "Meta-Analysis II."

[3/] The authors of Meta-Analysis I were Gene V Glass and Mary Lee Smith; when Meta-Analysis I is discussed in this critique, its authors are referred to as "Glass and Smith." The authors of Meta-Analysis II were Mary Lee Smith and Gene V Glass; conversely, when Meta-Analysis II is discussed, its authors are referred to as "Smith and Glass."

The principal investigator of the Class Size and Instruction Project was Leonard S. Cahen; he was assisted by Nikola N. Filby. Cahen is currently affiliated with Arizona State University and Filby with the Far West Laboratory. The Project was federally funded, primarily through a grant to the Far West Laboratory from the National Institute of Education. Additional support was provided through a NIE grant awarded to the Visiting Scholars Program, Center for the Study of Evaluation at UCLA.

The Glass and Smith meta-analyses, especially the meta-analysis on class size and pupil achievement, have received wide publicity. For example, the *New York Times* included an article on the Class Size and Instruction Project in April 1979. [19] Feature articles on Meta-Analysis I appeared in the January-February 1979 *Educational Evaluation and Policy Analysis* (authored by Glass and Smith) [14], the March 1979 *Phi Delta Kappan* (authored by Cahen and Filby) [5], and the National Education Association's April-May 1979 issue of *Today's Education* (authored by Glass, Cahen, Smith, and Filby). [17]

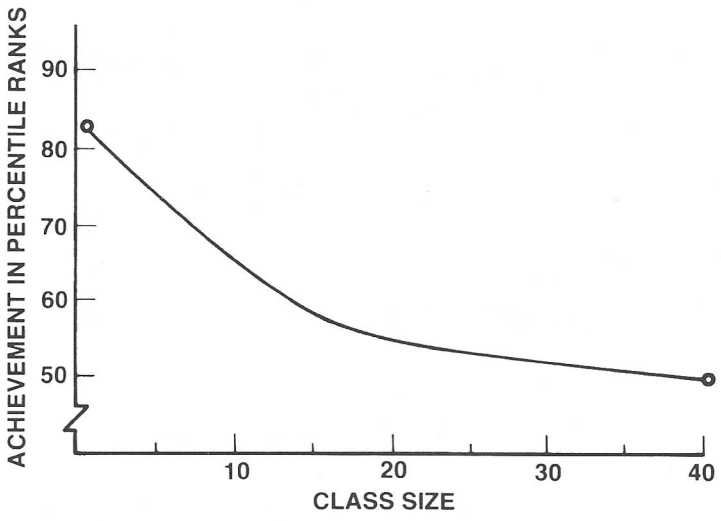
OVERVIEW

Basic Finding of the Meta-Analysis on Class Size and Pupil Achievement: What It Purports to Say and What It Actually Says

In their first meta-analysis on the effect of class size and pupil achievement, authors Gene Glass and Mary Smith stated that the best representation of their findings is shown in the graph below. This graph has since been widely published and presented by many as conclusive evidence that smaller classes result in increased pupil achievement.

The following brief discussion is presented as an overview of the central findings obtained through meta-analysis (which Glass described as an "analysis of analyses") and to provide the basis for further examination of the findings and interpretations of their analyses of class size research.

FIGURE 1.--Relationship Between Achievement and Class-Size. (Data Integrated Across Approximately 100 Comparisons from Studies Exercising Good Experimental Control.)



SOURCE: Gene V Glass and Mary Lee Smith. *Meta-Analysis of Research on the Relationship of Class-Size and Achievement.* San Francisco, California: Far West Laboratory for Educational Research and Development, September 1978, p. vi.

What is Figure 1? It is a graph which authors Glass and Smith said "form[s] the basis of our conclusion about how class-size is related to academic achievement." [16:v]

What is its importance? This graph has been widely presented as conclusive proof that pupils in smaller classes achieve more than those in larger classes. It has been used to assert that class size reduction should be a major educational goal and that ways must be found to finance smaller classes.

What is it based on? This graph is based on only 14 class size studies selected by the authors as being "well-controlled" (i.e., pupils randomly assigned to classes). Seventy-three percent of the comparisons used in the graph came from four studies. Eight of these 14 studies dealt with instructional conditions not typical to public schools (i.e., college classes or tutoring). The graph is *not* based on the total number of 76 studies and 725 comparisons the authors collected and used to justify the claimed superiority of their meta-analysis findings over all previous class size research analyses.

What does the graph measure? It is an *attempt* to "integrate" or statistically combine the effects of class size on achievement, "spanning the full range of class-sizes from individual tutorials to huge lectures" [16:32], across all subjects, among all grade levels (including college), and for pupils of all ability levels included in class size studies. The authors stated that it is the result of "using complex methods of regression analysis" that integrated many class size comparisons "into a single curve showing the relationship between class-size and achievement in general." [16:v] This process they term "meta-analysis." In simple, straightforward terms, it is a kind of statistical treatment in which the authors assigned numeric scores to the findings of certain class size studies dealing with the effects of class size on pupil achievement so they could present the results of the studies in a single curve.

What is its scale? It is presented in terms of "achievement in percentile ranks." However, these "percentile ranks" are merely highly generalized *hypothetical* metrics constructed theoretically from a conglomerate of achievement criteria, in various subjects, at various grade levels, and for pupils of various abilities. Although they appear to be similar to percentile ranks for specific standardized tests, these "percentile ranks" are *not* related to the range of scores on any standardized test of academic achievement in existence.

How do the authors describe the techniques used? The authors at one point described the analytical procedures they used in constructing the graph as "sophisticated methods of data analysis." [17:43] However, they later said while discussing the same methods they had used in constructing the graph that "the entire business was clumsy and inelegant." [30:26]

What does the graph show for the practical mid-range of class size? It indicates that even in the wide range of 20 to 40 pupils, class size makes *little difference* in pupil achievement. This is similar to the conclusion reached by most other reviewers of class size research; however, most others would be reluctant to agree to such a wide mid-range as one in which the research would indicate that class size makes little difference for most pupils.

Although viewed in another way, the graph indicates that class size could be doubled from 20 to 40 pupils without significant effect on pupil achievement, virtually all educators and laypersons would seriously question such an implication from the findings of Meta-Analysis I.

What does the graph show for very small classes? The authors stated that "the major benefits from reduced class-size are obtained as size is reduced below 20 pupils." [16:v] However, a substantial proportion of the comparisons used to influence this part of the graph dealt with extremely small instructional arrangements, such as one-to-one tutorial arrangements and classes of 2 to 5 pupils.

What does it show about pupils of different ability, in different grade levels, or in different subjects? According to the authors, neither "grade level, nor subject taught, nor ability of pupils" altered the basic graph of the relationship between class size and achievement. [16:v] Elsewhere they stated that the class size and pupil achievement effect was not influenced by "'source of data,' 'subject taught,' 'duration of instruction,' 'pupil IQ,' and 'type of achievement measure.'" [16:38]

What do the authors claim the curve shows? The authors stated that "we discovered many of the reasons why previous research reviewers lost their way in the forest of data and failed to find a defensible generalization [in favor of smaller classes]." [16:iv-v] The authors stated that meta-analysis permitted them to offer "bold generalizations" where previous class size reviewers were able to make only "timid qualifications." [16:iv]

What does the curve actually do? The meta-analysis procedure actually obscures and obliterates many of the helpful clues that are present in the existing class size research, such as those pertaining to pupils of varying abilities, various subjects taught, or various grade levels. In effect, the meta-analysis procedure statistically homogenized the findings of research studies in such a way as to lose helpful information from studies containing subtle but important clues. Even the authors themselves admitted that "in a very real sense, what will be done [in their meta-analysis] for the sake of arriving at general conclusions places the reader in benign jeopardy of losing qualitative and personal familiarity with the research." [16:4]

But more importantly, in attempting to integrate class size research "into a single curve" showing that smaller classes in themselves improve pupil achievement, the meta-analysis report places undue importance on the role of *general* class size reduction in the educational process.

What does it fail to do? The graph and related findings of the meta-analysis fail to provide information or guidelines that will be helpful to school officials, educators, and others in making class size decisions affecting pupils with specific abilities, in specific grades, or in specific subject areas. Instead, all that is offered is the "bold generalization" that smaller is better. And this is qualified by the admonition that classes must be *very small* to be much better.

SUMMARY OF META-ANALYSIS I

As Gene Glass (1976) defined the term, "meta-analysis" refers to "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings." [13:3] Thus, meta-analysis is a method of research integration; it is not original, experimental research. Glass added that "it connotes a rigorous alternative to the casual, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding research literature." [13:3] By using meta-analysis, Glass and Smith reported that they "discovered many of the reasons why previous [class size] research reviewers lost their way in the forest of data and failed to find a defensible generalization [in favor of smaller classes]." [16:iv-v] They noted the following problems with previous reviews of the class size literature:

> (1) literature searches were haphazard and often overly selective; dissertations were avoided, as a rule, and few reviewers sought out large archives of pertinent data; (2) reviews were typically narrative and discursive; the multiplicity of findings cannot be absorbed without quantitative methods of reviewing; (3) reviewers that attempted quantitative integration of findings made several mistakes: (a) they used crude classifications of class-sizes; (b) they took "statistical significance" of differences far too seriously; and (c) they lacked sufficiently sophisticated techniques of integrating results. [16:1-2]

As a result, researchers using traditional methods of analysis offered "timid qualifications . . . where bold generalizations were possible," the authors concluded. [16:iv] In contrast to the "inconclusive" results found in most traditional reviews of the class size literature, the "bold generalization" advanced in *Meta-Analysis of Research on the Relationship of Class-Size and Achievement* was that the class size research "established clearly that reduced class-size can be expected to produce increased academic achievement." [16:iv]

Glass and Smith collected 76 studies on the relationship between class size and pupil achievement, which they contended "exceeds by 50 percent the most extensive reviews published to date." [16:3][1/] These 76 studies yielded 725 comparisons between "smaller" and "larger" classes (which they termed "deltas") and were based on nearly 900,000 pupils of all ages and aptitudes and taking all types of subjects, according to the authors. These 725 comparisons ("deltas") then were integrated into a single curve showing the relationship between class size and pupil achievement. "This curve," stated Glass and Smith, "revealed a definite inverse relationship between class-size and pupil learning." [16:v] Moreover, when other circumstances were analyzed, "virtually none . . . altered the basic relationship; not grade level, nor subject taught, nor ability of pupils," they concluded. However, this curve did not appear in the final published edition of Meta-Analysis I. [16:v]

The only factor that influenced this relationship, the authors stated, was whether or not the original research studies in the data base were experimentally "well-controlled." The basis for their final conclusions of Meta-Analysis I rests with 110 comparisons from 14 studies that they considered to have been well-controlled. The curve derived from these "well-controlled" studies, according to the authors, is "the most accurate representation" of the class size-achievement relationship. [16:v] This curve appeared in the final version of Meta-Analysis I published in September 1978; it is reproduced as Figure 1 on page 6 of this critique.

This relationship is presented in terms of percentile ranks which the authors implied were related to those on nationally standardized achievement tests. A pupil who would score at the 83rd percentile on a national test when taught *individually* (tutorially), concluded the authors, would score at about the 50th percentile when taught in a class of 40. However, they found only a *six* percentile rank difference in favor of pupils taught in a class of 20 compared to those taught in a class of 40. "The major benefits from reduced class-size," Glass and Smith declared, "are obtained as size is reduced below 20 pupils." [16:v]

[1/] Although Glass and Smith stated that 77 studies comprised the total data base of Meta-Analysis I, the bibliography and the data listing actually contain 76 separate studies. Two of the entries reported data dealing with the same class size experiment. (See Study ID Number 52 in Appendix A on page 69.)

"Class size" was defined as the ratio of pupils to instructors. For example, a class size of 30 could have been one instructor with a group of 30 pupils, two instructors with a group of 60 pupils, or three instructors with a group of 90 pupils. Tabulations of class size effects were presented in terms of the relative differences found in comparing the effects of small vs. large classes termed "deltas" rather than the number of studies analyzed. Thus, a "small" class in one study could be a "large" class in another study. The data base for all 725 comparisons included the following variables:

- Year of the study--About half of the comparisons were drawn from studies before 1950, and half, from studies appearing between 1950 and 1978. [16:21]
- Publication source--Two-thirds of the total 725 comparisons were taken from journals. Another 18.9 percent of the deltas were found in theses and unpublished reports; 15.7 percent of all comparisons were taken from books. [16:23]
- Subject of instruction--Approximately half of the comparisons came from elementary school classes with "all subjects combined." Nearly 20 percent were taken from language classes and slightly more than 10 percent from mathematics classes. [16:24]
- Hours of instruction--This distribution showed modes of 50, 180, and 900 hours. Instruction ranged from one hour to 9,000 hours. [16:22, 25]
- Pupil age--About 40 percent of the comparisons included pupils from ages 5 to 10; about 60 percent included comparisons from pupils 11 and older, including college-age students. [16:26]
- Pupil IQ--A breakdown of this variable was not presented in the final version of Meta-Analysis I, but data from an earlier version of Meta-Analysis I showed that 59.0 percent of the pupils involved had an "average" IQ ($90 < IQ < 110$), 21.0 percent had a "high" IQ ($IQ > 110$), 7.6 percent had a "low" IQ ($IQ < 90$), and the IQ of pupils in 12.4 percent of the comparisons was "unknown." According to Glass and Smith, "the 'average' category is over-represented partly because a study that used a heterogeneous group of pupils and did not report findings separately for different IQ levels was coded as 'average.'" [15:33-34]

- Pupil assignment to "small" and "large" classes--The assignment of pupils to classes was "uncontrolled" in half of all comparisons, "matched" in one-third of the comparisons, and "random" in 15 percent of the deltas. [16:28]
- Achievement measure--Some type of nonstandardized test ("*ad hoc* measure") was used in 56.1 percent of the comparisons and some type of standardized test in 43.9 percent of the deltas. [16:30]

Of the 725 comparisons, which included studies "from individual tutorials to huge lectures," only 60 percent favored smaller classes in achievement. Furthermore, Glass and Smith stated that it is reasonable to suspect that the odds of observing a comparison favoring smaller classes in achievement in a "typical" class size of 15 to 40 pupils are perhaps as low as 55 percent to 45 percent. Given these odds, "one needs not wonder why narrative reviews of a dozen or two studies produced little but confusion," they said. [16:32]

The major conclusions reached by Glass and Smith in Meta-Analysis I were:

1. When the data were taken into account, the authors argued that: "A clear and strong relationship between class-size and achievement has emerged." [16:45] They added that "there is little doubt that, other things equal, more is learned in smaller classes." [16:46]
2. Class size and achievement effects were "consistently stronger in the secondary grades than in the elementary grades." [16:40]
3. Major differences in the relationship between class size and achievement were found depending on whether or not pupil assignment was random or uncontrolled. A curve based on random assignment of pupils to classes in 14 studies was presented as "probably the best representation of the class-size and achievement relationship." [16:43]
4. The relationship between class size and achievement was not affected by pupil IQ or different school subjects. [16:45]

SUMMARY OF META-ANALYSIS II

Published 10 months after the first meta-analysis on class size and achievement, *Relationship of Class-Size to Classroom Processes, Teacher Satisfaction and Pupil Affect: A Meta-Analysis* examined the effects of class size on nonachievement measures--classroom processes and the learning environment, student attitudes and behavior, and teacher satisfaction. Smith and Glass reported that Meta-Analysis II should be considered a "companion piece" to the first meta-analysis on class size and achievement, since the same procedures and methodology were used in both studies. The statistical techniques, however, were modified in Meta-Analysis II "because of the intervening development of improved techniques." [30:5]

In this review, Smith and Glass apparently used 60 studies which produced a total of 371 comparisons ("deltas").[1/] The data base for these 371 comparisons included the following variables:

- Year of the study--Sixty percent of the comparisons were drawn from studies dating from 1925 to 1969, and 40 percent from studies appearing between 1970 and 1978. One-fourth of all deltas came from studies published between 1975 and 1978. [30:10]
- Publication source--More than 60 percent of these comparisons were taken from theses and unpublished reports. Another 20.2 percent of the deltas were found in journals, and 14.0 percent, in books. [30:12]

1/ Although Smith and Glass stated that 59 studies comprised the data base of Meta-Analysis II [30:9], the bibliography and the data listing contain 60 studies.

- Subject of instruction--Almost 60 percent of the comparisons came from all-day self-contained classrooms, coded as "all subjects." Eleven percent were taken from social studies classes. Only 16 comparisons (4.3 percent of the total) were drawn from classes in reading, math, language, English, or writing. [30:12]
- Hours of instruction--This distribution showed a mean of over 450 hours. Instruction ranged from one hour to 6,000 hours. [30:11]
- Pupil/instructor ratio--As in Meta-Analysis I, P/I ratios were calculated by dividing the number of pupils in a group by the number of teachers teaching that group. The P/I ratio for "small" classes ranged from size 1 to size 78, with a mean of 20. The P/I ratio for "large" classes ranged from size 4 to size 189, with a mean of 45. [30:14]
- Pupil age--The mean age of pupils in the data base was 13, with a range of 5 to 22. [30:15]
- Pupil IQ--Smith and Glass noted that "if no information was provided by the researcher pupil ability was estimated as average. This was the usual situation, so that there was not sufficient variation to detect different class-size effects for different levels of pupil ability." [30:15]
- Pupil assignment to "small" and "large" classes--The assignment of pupils to classes was "uncontrolled" in 61 percent of all comparisons, "random" in 16 percent of the deltas, and "matched" in six percent of the comparisons. Seventeen percent of the comparisons came from "repeated measures" studies. [30:15]

Various nonachievement effects were separated into three categories: (1) affective effects on pupils, which contained 31 specific variables; (2) effects on teachers, which included seven specific components; and (3) effects on the instructional environments and processes, which contained 33 specific factors. [30:21-22]

The major conclusions reached in Meta-Analysis II were:

1. For the data as a whole, "these findings indicate that there is a beneficial effect on the general quality of the educational environment resulting from decreasing class size." [30:27] How these findings relate to pupils was discussed in the following quotation:

> Suppose that the typical level of non-achievement benefits experienced by the average pupil in a class of 30 pupils is set equal to the 50th percentile. The results . . . indicate that

> if this pupil were placed in a class of size 20, he would experience non-achievement benefits superior to 58 percent of the pupils who are taught in classes of size 30. In a class of 10 pupils, he will benefit more than 70 percent of the pupils in classes of 30, though he started out at the median (50th percentile) of such classes. On the other hand, increasing his class from 30 to 40 pupils would result in a decline in non-achievement benefits; 55 percent of the pupils in classes of 30 pupils would now experience greater benefits than he. In a class of 60 pupils, this hypothetical average student would gain benefits exceeding only 38 percent of the pupils in classes of 30 pupils. Even at this coarsest level of aggregation, class-size does make a difference. [30:27, 30]

2. The relationship between class size and nonachievement effects reported in Meta-Analysis II was more pronounced than the relationship between class size and achievement reported in Meta-Analysis I. [30:31]
3. Class size had a "substantial effect" on teachers. "The affective effects of class-size on pupils are positive but not as dramatic as the effects on teachers," Smith and Glass declared. [30:33] "The effect of varying class size on instructional processes and environments is the same as the effect on pupil affect." [30:33]
4. The effect of class size was greatest for pupils 12 years or younger, less for pupils from age 13 to 17, and least for pupils age 18 or older. [30:34]
5. Class size effects were *more* pronounced with comparisons from *uncontrolled* studies than for randomized or "well-controlled" studies. [30:34] However, unlike Meta-Analysis I, the final conclusions reached in Meta-Analysis II were based on comparisons from *all* studies, not just from the "well-controlled" studies. Nor were the conclusions reached in Meta-Analysis II based only on the uncontrolled studies, even though these class size effects were found to be more pronounced.
6. The effect of class size was greater in studies published from 1925 to 1968 than from 1969 to 1978. [30:38]
7. When classified by publication source, the effect of class size was greatest in unpublished papers, followed by books, journal articles, and theses. [30:38]

CRITIQUE OF THE TWO META-ANALYSES ON CLASS SIZE

After a careful examination of the two Glass and Smith meta-analyses and of subsequent interpretations and discussions of them, five major areas were identified that merited attention in this critique. The remainder of this report presents the five points that should be considered when class size issues are being discussed or studied in the context of the Glass and Smith meta-analyses.

Point 1--The method precludes indentification of meaningful clues contained in class size research.

* * * * * *

Point 2--The major findings rely on only a few studies, the methodology is inconsistently used, and the conclusions drawn are overgeneralized.

* * * * * *

Point 3--The interpretations of the findings are often contradictory.

* * * * * *

Point 4--The conclusions as presented only confuse the class size issue and unjustifiably encourage general class size reductions.

* * * * * *

Point 5--"Bold generalizations" create doubt about the need for further research.

POINT 1
Method Precludes Identification of Meaningful Clues Contained in Class Size Research

> The methodology used in the meta-analyses, by its very nature, smoothes out and obliterates important distinctions among variables contained in the class size research. Thus, important findings are lost in the process. The meta-analyses give the impression of accurate measurement through sophisticated statistical analysis, yet these measures are actually far too insensitive to identify many of the important relationships pertaining to class size.

Those familiar with the literature on class size know how difficult it is to draw absolute, unequivocal conclusions from the highly complex and interrelated factors relating to class size. This is why most previous class size reviewers deliberately have formulated cautious conclusions relating to the effect of class size on achievement and nonachievement measures. A methodology that would allow researchers to state unequivocally, once and for all, that smaller classes are better, regardless of any other factors, would be a monumental breakthrough in the field of educational research.

The two meta-analyses on class size have been advanced as just that kind of breakthrough. The stated purpose of these reports was to examine the class size research *quantitatively*, where others have analyzed it *qualitatively*. Because so-called quantitative methods termed "meta-analysis" were employed, it was assumed that the results were superior to all prior attempts to analyze the class size research.

In the summary of Meta-Analysis II, Smith and Glass stated that the relationships between class size and cognitive and noncognitive outcomes "have not in the past been apparent because of an inability to deal with either the class sizes or

the effects *precisely* and *quantitatively*. Using meta-analysis permits us to unravel the complexity and reveal the effects of class-size." [emphasis added] [30:ii] However, in examining the methods used and the findings derived from their research, certain factors should be clearly recognized.

While the meta-analyses convey the impression that *precise* measures of class size effects were provided, the methodology used in these reports actually obliterates important distinctions to be found in class size studies and, in so doing, distorts the findings of the existing research. Because the many class size effects contained in the literature were integrated to produce a single curve showing class size-achievement effects, the meta-analyses homogenized the data to such an extent that it became virtually impossible for the authors to conclude anything meaningful. Even members of the Class Size and Instruction Project admitted this in a later publication. "The meta-analysis technique," they said, "requires combining data from many studies and therefore does not allow fine distinctions among process variables." [10:3] "Bold generalizations" were provided in these meta-analyses, therefore, only because the methodology can provide little else. Such "bold generalizations" are of little value if meaningful implications and distinctions that could otherwise be drawn from the class size research are buried in the process.

While the meta-analyses state that a *quantitative* measure of class size effects was provided, the use of apparently sophisticated numeric measures gives the impression of accurate measurement, when in fact, these measures are far too insensitive to extract many of the subtle relationships relating to factors important to making class size decisions.

By dismissing the importance of statistical significance, the authors disregarded the previous experience and cautions of many class size investigators who conducted the original studies. It is difficult to understand why the authors said that previous analysts frequently "took 'statistical significance' of differences far too seriously." [16:2] However, Glass and Smith themselves stated that the "odds" of observing a finding that smaller classes are superior to larger classes in a typical classroom setting is "perhaps as low as 55% to 45%." [16:32]

It is equally difficult to understand why the authors at first described the analytical procedures they used in Meta-Analysis I as "sophisticated methods of data analysis" [17:43] and then later said in reference to their own use of these same techniques that "the entire business was clumsy and inelegant." [30:26]

The major conclusion reached in the meta-analysis on class size and achievement was:

> As class-size increases, achievement decreases. A pupil, who would score at about the 83rd percentile on a national test when taught individually, would score at about the 50th percentile when taught in a class of 40 pupils. The difference in being taught in a class of 20 versus a class of 40 is an advantage of 6 percentile ranks. The major benefits from reduced class-size are obtained as size is reduced below 20 pupils. [16:v]

The use of "percentiles" in this conclusion creates the illusion of precise measurement that is directly related to that of standardized tests, a type of measurement familiar to most educators and to many laypersons. Moreover, the specific reference to scores on "a national test" adds to the seeming precision of sophisticated measurement. However, it is important to understand what the "percentiles" mentioned in this conclusion and in many references to it really are. The "percentile ranks" used in the meta-analyses are merely highly generalized *hypothetical* metrics that are the results of arbitrarily combining various measures of achievement and nonachievement effects used in various class size studies, in various subject areas, at various grade levels, for pupils of various ability levels, and for various class sizes. Therefore, these "percentiles" are *not* and should *not* be related to any actual classroom situation or to any set of standardized test scores in existence.

Furthermore, while the use of "percentiles" and "percentile ranks" in the reports gives the impression of precise measurement that seemingly would be valuable in making class size decisions, in reality, it offers no helpful clues concerning efficient class sizes for specific groups of pupils, in specific subject areas, or at specific grade levels. The findings imply that the only thing *any* school district needs do to raise the achievement level of *any* group of 40 pupils, at *any* ability level, in *any* subject level, or at *any* grade level on *any* nationally standardized test from an average score at the 50th percentile to an average of the 56th percentile would be to divide the class into two classes of 20. And if the school district wanted to raise the average score for the 40 pupils to the 83rd percentile, all that would be necessary would be to provide an individual tutor for each of the 40 pupils. In view of this, the major conclusion, it is evident that the process of meta-analysis actually has obliterated any important distinctions among variables that may be contained in the existing body of class size research.

POINT 2

Relies on Few Studies, Methodology Inconsistently Used, and Conclusions Overgeneralized

The methodology used in the meta-analyses purportedly was designed to correct alleged deficiencies in previous reviews of the class size research, in particular, the imprecise reliance on a few studies from which timid conclusions had been drawn. Actually, the major findings of the meta-analyses rest on *fewer* studies than previous reviews, not more; the methodology was applied inconsistently from one meta-analysis to the other; and the conclusions reached in these meta-analyses were greatly overgeneralized.

COMPOSITION OF THE DATA BASE USED IN META-ANALYSIS I

Since the general finding on the relationship of class size and pupil achievement was based on only 14 of the total 76 studies in the data base, how "extensive" actually was Meta-Analysis I compared to the previous research? Meta-Analysis I had 14 studies as the basis for its "bold generalizations" compared with Lindbloom's 1970 review (13 studies on class size and achievement), the 1975 summary published by the New England School Development Council (18 studies on class size and achievement), and the 1978 ERS summary of research on class size (35 studies on class size and achievement). With their reliance on only 14 "well-controlled" studies for their final results, it is curious that the authors would challenge the scope of other class size reviews when the data base on which their conclusions were drawn actually was less than the data base of other studies. In addition, *this data base of 14 studies actually is comprised of only six studies that relate to class sizes and conditions normally found in the elementary and secondary grades.*

Readers who wish to take a closer look at the 110 comparisons in the 14 "well-controlled" studies of Meta-Analysis I cannot do so from the material

contained in the published report. This is despite the inclusion of the study's data listing in an appendix and a description of items reported on coding sheets that were used in processing the data. The problem with interpreting the 12 pages of coded data is that there is no key to this listing in Meta-Analysis I with which to link the study with its comparisons. Given this omission, the inclusion of the data listing is of little value. It also tends to be misleading, since readers might assume that the raw data used for the meta-analysis was given for close examination by interested persons, when, in fact, this is not possible.

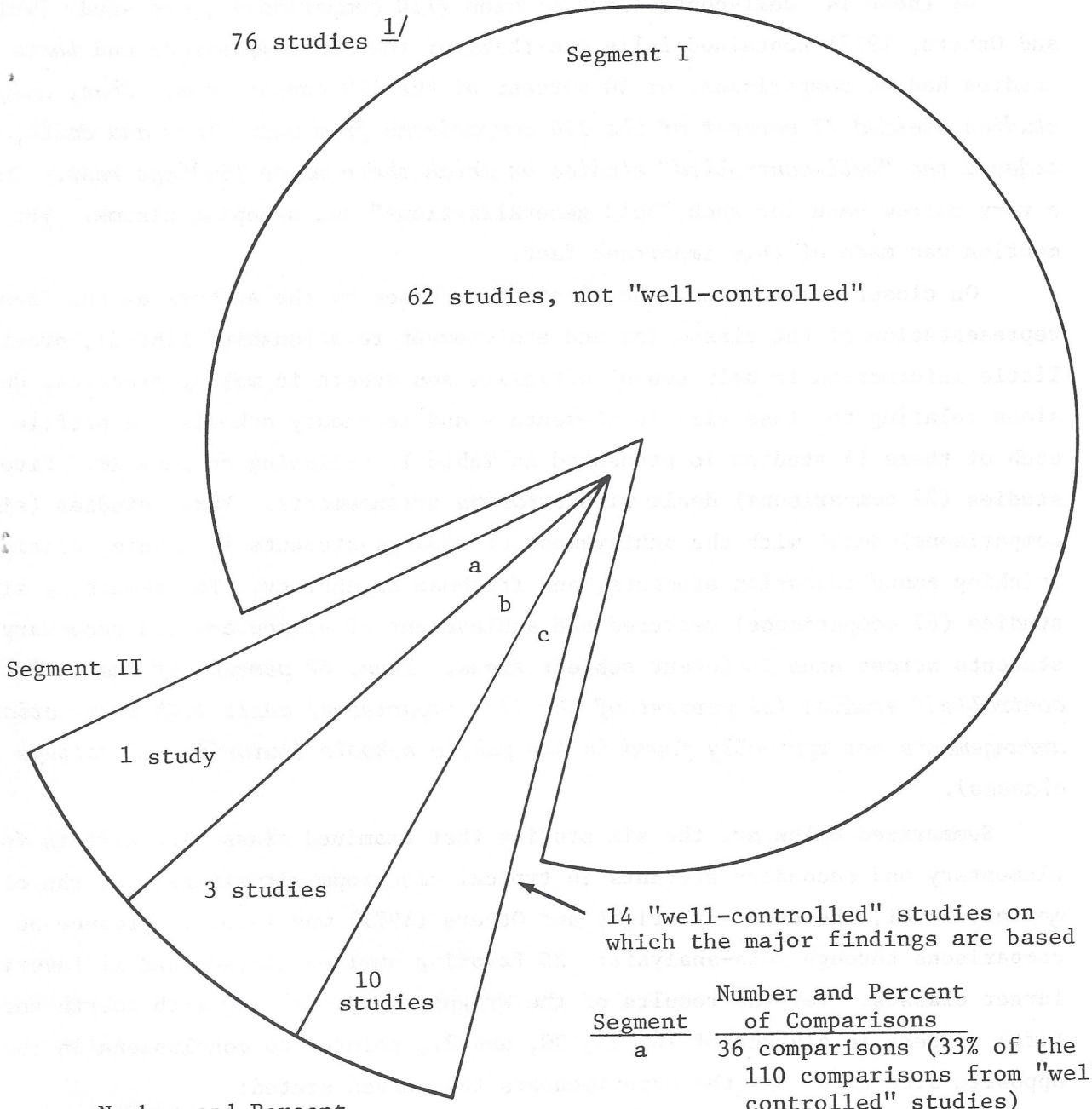
On page 54 of Meta-Analysis I, Glass and Smith stated that "the key to decoding the variables [in the data listing] appears in Table 3.1 in the section of the report on Methods." Yet there is no Table 3.1 in Meta-Analysis I, which was published in September 1978; rather, Table 3.1 appeared on pages 14 and 15 of a May 1978 edition of Meta-Analysis I. It seems that the authors intended to put this important table in their final published report, but that proper care was not taken to see that it actually was included.[1/]

From the earlier unpublished edition of Meta-Analysis I, it is possible to interpret the coding of the data listing in each of the 76 studies and 725 comparisons in the published edition. On examining the entries in the data listing with the key to decoding the variables, ERS found a number of apparent coding errors. A discussion of these errors would be too involved to be included in this critique. In drawing conclusions from the data listing, ERS used the data just as they appeared in the listing, since the meta-analysis report was based on such data.

From the earlier edition of Meta-Analysis I, it is possible to discern which studies comprised the narrower data base of 14 "well-controlled" studies (110 comparisons) on which the final analysis of Meta-Analysis I rests. These 14 studies, their ID numbers, and the number of comparisons in each study are listed in Appendix A, on page 69 of this critique. Presented in Figure 2 is a graphical representation of the studies and the number of comparisons used to make the final conclusions of Meta-Analysis I. *The 110 comparisons in "well-controlled" studies represent only 15 percent of the total 725 comparisons for all studies that Glass*

[1/] However, Smith and Glass did include in Meta-Analysis II as part of the data listing the key to decoding the variables. (See pages 63 and 64 in Meta-Analysis II.)

FIGURE 2.--Number of Studies and Number of Comparisons ("Deltas") in the Data Base of Meta-Analysis I



Segment	Number and Percent of Comparisons
I	615 comparisons (85% of all 725 comparisons)
II	110 comparisons (15% of all 725 comparisons)

N = 725 total comparisons

Segment	Number and Percent of Comparisons
a	36 comparisons (33% of the 110 comparisons from "well-controlled" studies)
b	44 comparisons (40% of the 110 comparisons from "well-controlled" studies)
c	30 comparisons (27% of the 110 comparisons from "well-controlled" studies)

N = 110 comparisons from "well-controlled studies

1/ See the footnote to Appendix A on page 70.

and Smith initially reviewed. The 14 "well-controlled" studies accounted for only 18 percent of the total data base of 76 studies.

Of these 14 "well-controlled" studies (110 comparisons), *one* study (Wright and Others, 1977) contained fully one-third of the 110 comparisons and *three* other studies had 44 comparisons, or 40 percent of the 110 comparisons. *Thus, only four studies yielded 73 percent of the 110 comparisons from what Glass and Smith considered the "well-controlled" studies on which their major findings rest.* This is a very narrow base for such "bold generalizations" and sweeping claims. Yet no mention was made of this important fact.

On closer examination, the 14 studies chosen by the authors as the "best representation of the class-size and achievement relationship" [16:43], provide little information to help school officials and others in making practical decisions relating to class size in elementary and secondary schools. A profile of each of these 14 studies is presented in Table 1, beginning on page 28. Five studies (37 comparisons) dealt with *tutoring* arrangements. Three studies (six comparisons) dealt with the achievement of *college* students in tennis, critical thinking among education students, and freshman psychology. The remaining six studies (67 comparisons) measured the achievement of elementary and secondary students across nine different subject areas. *Thus, 57 percent of the 14 "well-controlled" studies (39 percent of the 110 comparisons) dealt with instructional arrangements not typically found in the public schools (tutoring and college classes).*

Summarized below are the six studies that examined class size effects for elementary and secondary students in typical classroom situations. At the elementary level, the study by Wright and Others (1977) was found to produce 36 comparisons through meta-analysis: 25 favoring smaller classes and 11 favoring larger classes. Yet the results of the Wright study, dealing with fourth and fifth graders in classes of 16, 23, 30, and 37, pointed to conclusions in the opposite direction. As the experimenters themselves stated:

> Standardized tests that were used to measure academic achievement showed a statistically significant class size effect in one of the four areas assessed, Mathematics--Concepts. There were no significant differences found for the measures of Reading, Vocabulary or Mathematics--Problem Solving. Measures of the students' development in Art and Composition based on samples of the students' work showed no differences between class sizes. An argument against standardized tests was that performance in a creative endeavour such as Art or Composition would be more sensitive to class size effects than performance on a multiple choice test. This was not supported. Students showed no sign of development on the art measure over the two years. [32:123]

Meredith, Johnson, and Garcia-Quintana (1978) found that for the first graders studied in South Carolina, class size had no statistically significant effect on reading, mathematics, language, or overall achievement. Balow (1969) reported that class size positively affected reading achievement when pupils were in smaller classes (size 15) over two or more consecutive years. First grade was the year critical to reading achievement; by the third grade, class size alone was not the only factor in determining pupil achievement. Boys gained more from smaller classes than girls, but pupils of all abilities seemed to benefit from smaller classes.

At the secondary level, Smith (1974) found that smaller classes (size 28) effected significant improvement in the knowledge of writing skills and in writing performance of the high school juniors studied, but individualized instruction produced even greater effects. These improvements also were greater for low- and average-achieving students than for high-achievers. Low- and average-achievers who were given individualized instruction had higher levels of retention of writing knowledge and improvement than high-achievers. Previous small-class instruction produced no significant difference in retention of improved writing performance, but there was a retention in knowledge of writing skills. Students at all levels of achievement who had received small-class instruction for six weeks scored higher on a post-posttest objective test than students in large classes. Haskell (1964) found that class size made no significant difference in achievement in geometrical drawing for a sample of first year secondary students. Cram (1968) and Jeffs and Cram (1968) reported that class size made no significant difference in the achievement of students in two business classes, but that students in a government class of 24 had greater achievement gains than students in a larger class of 45.

At the college level, Verducci (1969) measured class size and achievement effects for students in a tennis class. The achievement measure was the ability to hit a tennis ball above a white line on a wall or backboard as many times as possible for 30 seconds. DeCecco (1964) studied class size and achievement effects in a freshman psychology course divided into "small control" groups (mean of 28 students), "small experimental" groups (mean of 28 students), and two "large experimental" groups (size 97 and 127). Bostrom (1968) examined the relationship between class size and critical thinking skills of 90 beginning students admitted to the College of Education at Arizona State University.

Of the *four* studies which accounted for nearly *three-fourths* of the comparisons from the 14 "well-controlled" studies used in the final analysis of Meta-Analysis I, two dealt with *tutoring*, one with traditional class sizes, and one with both traditional class sizes and individualized instruction. The one study on which a third of the 110 comparisons were based found a significant relationship between class size and pupil achievement in *only one of six subject areas examined.* In another study, pupil ability was an important determinant of the achievement effects.

Five of the 14 "well-controlled" studies (36 percent) dealt with tutoring arrangements. These five studies produced, through meta-analysis, 37 comparisons (34 percent of the 110 deltas from the 14 studies). Thirty-four of these 37 comparisons favored smaller classes.

Examination of these 14 "well-controlled" studies provides insight into the effects of meta-analysis. It shows how distinctions that could otherwise be observed among important variables in the class size research are obliterated in this use of the meta-analysis process. The technique used to integrate the findings of these studies is so insensitive to other achievement-related variables that it precludes detection of other important factors. Moreover, it seems that in some cases meta-analysis leads to conclusions that contradict the conclusions reached by the original experimenters themselves. For example, see the study (ID# 35) conducted by Wright and his associates who researched class size effects relating to fourth and fifth graders in Toronto. This study, which accounted for one-third of the 110 comparisons, found that class size made no difference in five of six subject areas examined. However, these findings were interpreted through meta-analysis in such a way that 25 of the 36 comparisons favored smaller classes. The study conducted by Meredith, Johnson, and Garcia-Quintana (ID# 49) found that class size had no statistically significant impact on achievement in reading, mathematics, language, or overall achievement. However, all four of the comparisons assigned to this study favored small classes. Thus, meta-analysis seemed to place the analysts in the position of supposedly knowing more about the research findings of specific studies than the original experimenters themselves.

The composition of this data base of 14 studies also raises questions about the prudence of the whole data selection procedure used in the meta-analyses. In the attempt to integrate more research on class size than had ever been done before, the authors included a number of class size studies that were irrelevant to typical school situations. Rather than being "overly selective" in the choice of class

size studies, a criticism that Glass and Smith made of previous class size reviews, the authors took the opposite approach in selecting studies for inclusion in the meta-analyses; eight of the 14 "well-controlled" studies selected were based either on tutoring arrangements or college classes.

In the final analysis, persons interested in learning how class size affects specific pupils in specific subject areas in group settings typically found in the public schools cannot find an answer here. When analyzed in this manner, the data base of 14 "well-controlled" studies is patently inadequate to offer clues for effective class size decisions. Tutoring studies and studies involving college students do not belong in a data base from which generalizations will be made relating to pupils in elementary and secondary schools. Contrary to what is concluded in Meta-Analysis I, these 14 "well-controlled" class size studies hardly represent the "best representation of the class-size and achievement relationship." Previous reviews of the class size research, which the authors termed "haphazard and often overly selective" [16:1], have actually covered the field more thoroughly and have provided far more helpful information for interpreting the class size research than Meta-Analysis I, despite its claim.

TABLE 1.--Profiles of the 14 Class Size Studies Classified as "Well-Controlled" in Meta-Analysis I Showing Summaries Based on the Original Studies and Some Classifications Contained in Meta-Analysis I

SUMMARY BASED ON ORIGINAL STUDY*						SOME CLASSIFICATIONS FROM META-ANALYSIS I**							
Study	Type of Study	Class Sizes	Grade Level	Pupil Ability	Subjects	Number of Deltas	Percent of 110 Deltas	Cumulative Percent of 110 Deltas	P/I Ratios Small	P/I Ratios Large	IQ	Number of Deltas Favoring Small	Number of Deltas Favoring Large
Wright and Others (1977) (Study ID #35)**	Class size	Smaller: 16, 23 Larger: 30, 37 (Variation of ± 2 allowed.)	4-5	Pupils were randomly assigned to classes, stratified by sex and a rating of academic performance.	Reading Comprehension; Vocabulary; Mathematics--Problem Solving; Mathematics--Concepts; Art; Composition	36	32.7%	32.7%	16, 23, 30	23, 30, 37	Average	25	11
Experimenters' Findings: Class size had a significant effect on only one of the six areas assessed (Mathematics--Concepts).													
Shaver and Nunn (1971) (Study ID #3)	Tutoring	1-to-1 and 1-to-3 instruction was compared with instruction in control groups (size unspecified).	4, 7, 10	Under-achievers (classified as those students whose reading and writing scores on the *Sequential Tests of Educational Progress* were lower than scores predicted from the *California Test of Mental Maturity*.)	Reading and Writing	18	16.4%	49.1%	1, 3	3, 25	High	16	2
Experimenters' Findings: Students at all three grade levels who were tutored in 1-to-1 or 1-to-3 arrangements had greater end-of-year standardized test scores than students in control groups. This finding was sustained two years later for seventh and tenth graders who had been tutored. There were significantly greater frequencies of tutored students at all three grade levels who reached their predicted potential or better, as compared to students in control groups. This difference also was present two years later.													
Smith, D.I. (1974) (Study ID #77)	Class size and individualized instruction	Small: 28 Experimental: 28 Large: 37	11	Low, average, and high achievers	Writing	18	16.4%	65.5%	1, 14	14, 30	Low, Average, High	15	3
Experimenter's Findings: Small classes can effect significant improvement in the knowledge of writing skills and in writing performance, but individualized instruction can produce even greater effects. These improvements were greater for low- and average-achievers than for high-achievers. In terms of retention, low- and average-achievers who were given individualized instruction had higher levels of retention of writing knowledge and improvement than high-achievers. In terms of retention, previous small-class instruction produced no significant difference in improved writing performance, but there was a retention in knowledge of writing skills. Students at all levels of achievement who had received small-class instruction for six weeks scored higher on a post-posttest objective test than students in large classes.													

Ronshausen (1975) (Study ID #55)	Tutoring	1-to-1	K-2	Under-achievers, pupils of low ability, or pupils with low readiness test scores.	Mathematics	8	7.3%	72.8%	1	30, 33	Average	8	0
Experimenter's Findings: Tutoring produced increases in achievement test scores "which have convinced local school officials that peer tutoring is beneficial to the tutees" in three experimental projects in Kentucky, Indiana, and California. Although each of the three projects produced achievement test scores which indicated the value of tutoring, these data were not presented in this paper.												
Moody and Others (1973) (Study ID #9)	Tutoring	Smaller: 1, 2, 5 Larger: 23	4	Not controlled	Mathematics	6	5.5%	78.3%	1, 2, 5	2, 5, 23	Average	6	0
Experimenters' Findings: Test scores given immediately after a lesson were greater for pupils in classes of 1, 2, and 5 than those in the class of 23. One-to-one instruction was found to be significantly superior than 1-to-5 instruction.												
Meredith, Johnson, and Garcia-Quintana (1978) (Study ID #49)	Class Size	Small: 21 Large: 29	1	Five of the 25 participating districts first attempted to stratify pupils by race and/or sex and/or ability before using a random procedure to select pupils to classes. The remaining 20 districts did not stratify pupils according to these factors before random selection was made.	Reading, Mathematics, Language, and Overall Achievement	4	3.6%	81.9%	20	27, 28	Average	4	0
Experimenters' Findings: Class size had no statistically significant effect on reading, mathematics, language, or overall achievement of the first graders involved in the project.												

*These summaries were made by ERS using copies of the original studies cited in Meta-Analysis I.

**There are 25 possible coded classifications in the data listing of Meta-Analysis I. Only those indicating the study identification number (column 1), the number of "deltas" (calculated by adding the number of rows for each ID number), pupil/instructor ratios for small and large classes (columns 9 and 13, respectively), IQ (column 15), and the number of deltas favoring small classes and large classes (columns 24 and 25, respectively) are shown here.

(continued)

TABLE 1 *(Continued)*

SUMMARY BASED ON ORIGINAL STUDY						SOME CLASSIFICATIONS FROM META-ANALYSIS I							
Study	Type of Study	Class Sizes	Grade Level	Pupil Ability	Subjects	Number of Deltas	Percent of 110 Deltas	Cumulative Percent of 110 Deltas	P/I Ratios Small	P/I Ratios Large	IQ	Number of Deltas Favoring Small	Number of Deltas Favoring Large
Haskell (1964) (Study ID #6)	Class Size	Small: 17 Large: 34, 35	Secondary, first year	Pupil intelligence was one of five factors held constant.	Geometrical Drawing	4	3.6%	85.5%	17	35	High	1	3
Experimenter's Findings: No significant differences in achievement were found between small and large groups in the first or second terms, but a significant effect at the .05 level was noted for the third term. "This is not sufficient evidence to accept, with any high degree of confidence, that size of class will make a significant difference to achievement in the subject."													
Bausell, Moody, and Walzl (1972) (Study ID #1)	Tutoring	1-to-1 instruction was compared with instruction in a regular classroom setting (size unspecified)	4-5	Low, medium, and high ability pupils were included in the study.	Mathematics	3	2.7%	88.2%	1	25	Low, Average, High	3	0
Experimenters' Findings: Tutoring produced significantly greater achievement scores than classroom instruction. Ability level was not affected by the instructional arrangement used.													
Cram (1968), Jeffs and Cram (1968) (Study ID #52)	Class Size	Control ("average"): 24, 26, 27 Experimental ("above average"): 45, 50, 52	9-12	Student assignment to classes was not stratified by student ability. However, *t*-test scores indicated no significant differences in mean intelligence scores between students in experimental and control groups.	Business, Government	3	2.7%	90.9%	24, 26 27	45, 50, 52	High	2	1
Experimenters' Findings: No achievement differences were noted in smaller or larger classes in either of two business courses offered (introduction to business and business law), but smaller classes in government showed higher achievement than larger classes.													

Verducci (1969) (Study ID #58)	Class Size	Small: 15 Medium: 37 Large: 60	College	Not controlled	Physical Education--hitting a tennis ball above a white line on a wall or backboard as many times as possible in 30 seconds.	3	2.7%	93.6%	15, 37	37, 60	High	3	0

Experimenter's Findings:
Students in a small class performed significantly higher than students in a large class. However, there was no significant difference in achievement between students in the small vs. the medium class or in the medium vs. the large class.

Balow (1969) (Study ID #16)	Class Size	Small: 15 Large: 30	1-4	Standardized test scores were analyzed according to IQ and other variables.	Reading	2	1.8%	95.4%	15	30	Average	2	0

Experimenter's Findings:
When students were in small classes for two or more consecutive years, class size positively affected pupil achievement. First grade was critical to reading achievement; by the third grade, class size alone was not the only factor in determining achievement. Boys gained more from smaller classes than girls. Reduced class size appeared to be beneficial to pupils at all levels of aptitude.

DeCecco (1964) (Study ID #8)	Class Size	"Small control": range of 22-35, mean of 28 "Small experimental": range of 18-34, mean of 28 "Large experimental": 97, 127	College freshman	Not controlled	Psychology	2	1.8%	97.2%	28	112	High	1	1

Experimenter's Findings:
Class size was not a significant variable in determining student achievement in large vs. small groups or in groups where instructional materials and assignments were coordinated vs. uncoordinated.

(continued)

TABLE 1 *(Continued)*

SUMMARY BASED ON ORIGINAL STUDY						SOME CLASSIFICATIONS FROM META-ANALYSIS I							
Study	Type of Study	Class Sizes	Grade Level	Pupil Ability	Subjects	Number of Deltas	Percent of 110 Deltas	Cumulative Percent of 110 Deltas	P/I Ratios Small	P/I Ratios Large	IQ	Number of Deltas Favoring Small	Number of Deltas Favoring Large
Ellson and Others (1965) (Study ID #61)	Tutoring	Unable to determine	K-1	Subjects included retarded children, slow readers, and unselected populations of K-1 school children	Reading	2	1.8%	99.0%	1	8	Low	1	1
Experimenters' Findings: This paper summarized 10 experiments involving 400 pupils in which the technique of programed tutoring was used. Results of several experiments found that programed tutoring was most successful when used as a supplement to and coordinated with regular classroom teaching.													
Bostrom (1969) (Study ID #73)	Class Size	Small: 15 Large: 45	College	Students of high, low, and average ability were included in the sample.	Critical thinking, education students	1	0.9%	99.9%	15	45	High	1	0
Experimenter's Findings: Class size had no significant effect on student achievement. However, high achieving students learned critical thinking skills more efficiently in large classes, while low achievers learned more efficiently in smaller classes. The achievement effects for average ability students were similar regardless of class size.													

COMPOSITION OF THE DATA BASE USED IN META-ANALYSIS II

The data base used in the second meta-analysis also was examined, although not in the same detail as Meta-Analysis I. Even so, this investigation yielded some important findings of which readers should be aware when they interpret the conclusions reached in the second report.

As in Meta-Analysis I, ERS found apparent coding errors in the data listing of Meta-Analysis II. However, in drawing conclusions from the data listing, ERS used the data just as they appeared in the listing of Meta-Analysis II, since the report was based on such data.

As in Meta-Analysis I, the selecting of data for inclusion in Meta-Analysis II was so broad that the data base contained many studies that did not belong in a report dealing with class size effects on elementary and secondary students. *Of the 60 studies examined in Meta-Analysis II, 12 (20 percent) dealt with students age 19 or older. This amounted to 43 of the 371 comparisons, or 12 percent.*

On the other hand, the data base of Meta-Analysis II contained only 16 comparisons (4.3 percent of all 371 deltas) that were drawn from classes specifically in reading, math, language, English, or writing.

On page 14 of Meta-Analysis II, the authors stated that the pupil/instructor ratios (their measure of "class size") in the 371 deltas analyzed ranged from 1 to 78 in small classes and from 4 to 189 in large classes. On examination, many of the comparisons contained in the data base of Meta-Analysis II were based on either very small or very large P/I ratios. Sixty-nine of the 371 deltas (19 percent) were based on a P/I ratio in small classes of 10 or less. Seventy-six of the 371 deltas (20 percent) were based on a P/I ratio in large classes of 50 or more. When these figures are adjusted to account for overlapping between these categories, *127 of the 371 comparisons (34 percent) were based on a P/I ratio in small classes of 10 or less or a P/I ratio in large classes of 50 or more.*

In Meta-Analysis II, 60 studies produced 371 class size comparisons, with each study, on the average, having about six comparisons. However, when the data base is examined beyond this average calculation, very few of these 60 studies substantially influenced the final results of the meta-analysis. In particular,

- 2 of the 60 studies accounted for 90 comparisons (an average of 45 deltas per study). Thus, these two studies (three percent of all studies) produced fully 24 percent of all comparisons.

- 6 of the 60 studies accounted for 159 comparisons (an average of 27 deltas per study). Thus, these six studies (10 percent of all studies) produced 43 percent of all comparisons.
- 37 of the 60 studies accounted for only 84 comparisons (an average of 2 deltas per study). Thus, these 37 studies (62 percent of all studies) produced only 23 percent of all comparisons.

As shown graphically in Figure 3, a situation exists in which the results of *two studies* had as much impact on the data base and the results of Meta-Analysis II as *37 studies*. No mention was made of this significant fact. To repeat a basic point, the authors' continued insistence that the past class size research was overly selective, and that their meta-analysis remedied this shortcoming, does not hold up under close scrutiny. Meta-Analysis II, like its predecessor, was based to a large extent on the findings of a few selected class size studies.

GENERALIZATIONS ON THE EFFECT OF CLASS SIZE ON NONACHIEVEMENT MEASURES

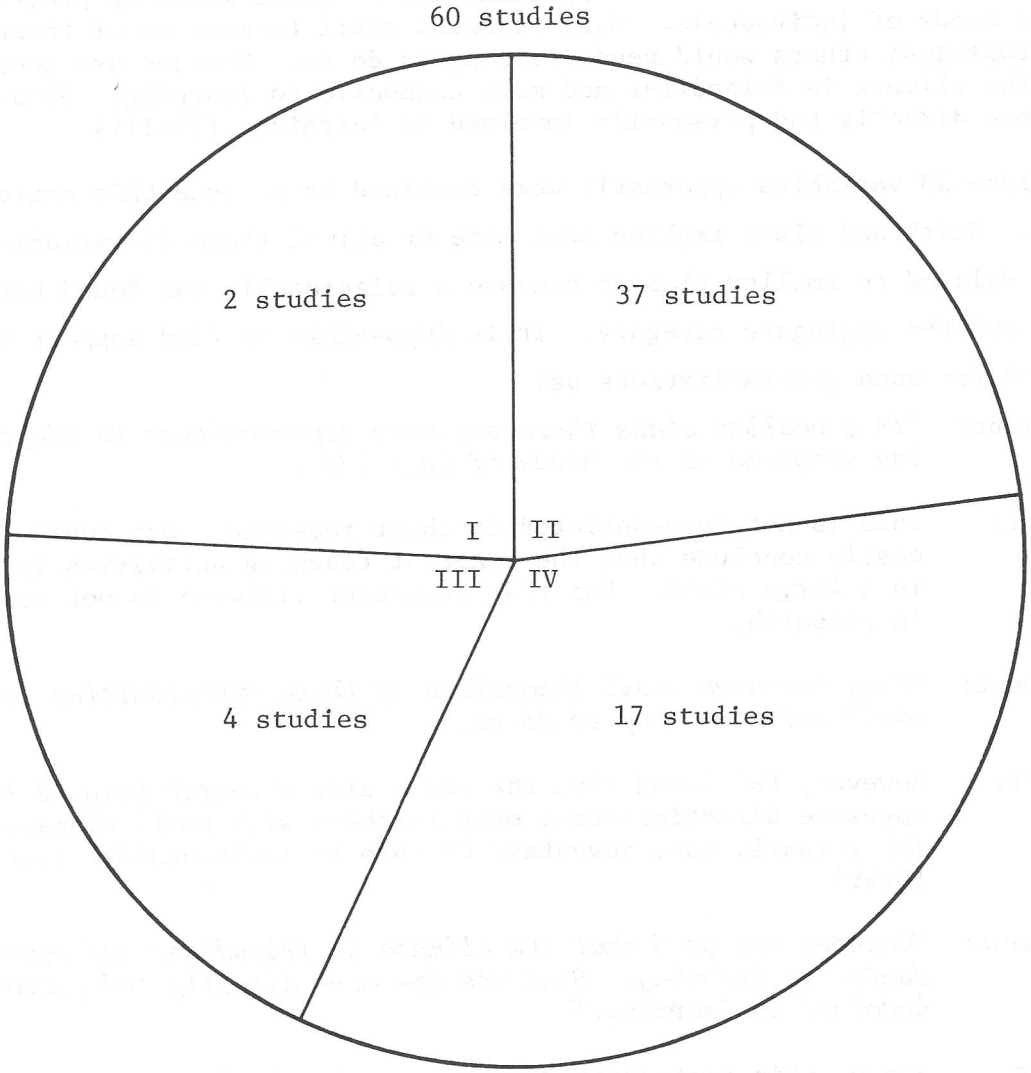
In the 1977 edition of *Review of Educational Research*, Glass described a potential hazard of the meta-analysis approach to data integration:

> Very likely, any investigator will find a large number of irresistible questions to try to answer in a meta-analysis. Even if he collects and analyzes over a thousand studies, he will still attempt to answer questions which will stretch the available data very thin. [11:362]

As the following discussion illustrates, the data in Meta-Analysis II were stretched very thin indeed when generalizations on the effect of class size on nonachievement measures were made.

Meta-Analysis II presented conclusions based on three categories of "quality education." Thirty-one separate variables were identified in the category "affective effects on pupils," 33 different factors made up "effects on instructional environments and processes," and seven variables constituted the category "effects on teachers." The total number of variables in these three groups (71) then were combined to produce the aggregate category called "quality of the educational environment." The effects of class size on each of these three groups and on the three categories taken as a whole then were discussed. From an analysis of these conclusions, it appears that the authors overgeneralized their

FIGURE 3.--Number of Studies and Number of Comparisons ("Deltas") in the Data Base of Meta-Analysis II



Segment		Number and Percent of Comparisons
I		90 comparisons (24% of all 371 comparisons)
II		84 comparisons (23% of all 371 comparisons)
III		69 comparisons (19% of all 371 comparisons)
IV		128 comparisons (34% of all 371 comparisons)
N	=	371 total comparisons

data to such an extent that the implications drawn from Meta-Analysis II are not supported even by their own research findings.

Generalization 1:

> Class size affects the quality of the classroom environment. In a smaller class there are more opportunities to adapt learning programs to the needs of individuals. Many teachers avail themselves of these opportunities; others would need training to do so. Chances are good that the climate is friendlier and more conducive to learning. Students are more directly and personally involved in learning. [30:iii]

Data from 33 variables apparently were combined to produce this aggregate conclusion. Smith and Glass implied that some or all of these 33 variables were positively related to smaller classes because a relationship was found between class size and the aggregate category. It is impossible to find support *in their own research* for such generalizations as:

Statement: *"In a smaller class there are more opportunities to adapt learning programs to the needs of individuals."*

Comment: This is not substantiated in their research. One could just as easily conclude that the range of teaching activities is greater in a large class. But that statement likewise is not supported in research.

Statement: *"Many teachers avail themselves of these opportunities; others would need training to do so."*

Comment: However, ERS found that the class size research pointed in the opposite direction--that many teachers with small classes do *not* actually take advantage of them to individualize instruction. [6:69]

Statement: *"Chances are good that the climate is friendlier and more conducive to learning. Students are more directly and personally involved in learning."*

Comment: Again, this contention may be true, but it is unsupported by any research evidence presented in their meta-analysis.

It is interesting to note that the results of the Far West Laboratory's 1980 class size study, *What Happens in Smaller Classes? A Summary Report of a Field Study*, do *not* support the contention that most teachers take advantage of smaller classes to individualize instruction. This study reported the results of what occurred when class sizes were deliberately reduced from 20 to 13 in two second grade classes in Virginia and from 35 to 22 in two second grade classes in California. Filby and her associates stated that "teachers were encouraged to take

advantage of the small class and thus put some energy into making changes." [10:15] Even so, few teachers individualized instruction:

> No teacher really individualized the curriculum. Groups of different sizes were the basic organizational unit. Some teachers provided supplementary individual work but to a limited degree. One teacher talked about preparing individual learning packets in math but did not do so. . . .
>
> Only one teacher tended to include activities that integrated different subject areas. There were few opportunities for student choice of activities. Teachers still tended to evaluate student work through brief written symbols and brief verbal comments, rather than more extended personal discussions of progress. [10:16]

* * * * * * * * * * * *

> For those who desire major changes in education in general or in a particular teacher, class size is not the whole answer. Clear differences exist between teachers in style and effectiveness. Small classes do not eliminate these differences or bring major changes in educational approach. Small classes do give the teacher a chance to implement an approach more fully and more effectively. [10:17]

Generalization 2:

> Class size affects pupils' attitudes, either as a function of better performance or contributing to it. In smaller classes, pupils have more interest in learning. Perhaps there is less distraction. There seems to be less apathy, friction, and frustration. [30:iii]

The effect of varying class size on pupil affect was the same as the effect on instructional processes and environments, said Smith and Glass. [30:33] But nowhere in the research did the authors provide evidence in support of the three qualifying statements in the passage above. Again, no evidence is given as to how specific conclusions like those illustrated in the qualifying statements above can be isolated from 31 factors that were lumped together for purposes of aggregation. Although these statements may strengthen the conclusion of Meta-Analysis II, they were not warranted by the research contained in the report.

Generalization 3:

> There was a substantial effect of varying class size on teachers. The difference in a teacher's workload, attitudes about students, morale, and general satisfaction varies from the 50th percentile in a class of 30 pupils to the 76th percentile in a class of 15. The difference in teacher effects in a class of 10 versus a class of 40 is 49 percentile ranks. Thus the truism is given empirical support: teachers feel better and feel they perform better in smaller classes. [30:33]

In the graphical representation of this relationship, it is interesting to note that the curve for pupil effects and instruction effects was based on 327 comparisons, yet the curve for teacher effects was based on only 30 comparisons. [30:32][1/] Said in another way, there were 11 times as many comparisons relating to pupil and instruction effects as to teacher effects. Yet no cautions were made about drawing inferences from these curves.

When the 31 comparisons that comprised the teacher effects variable were examined, 13 of these comparisons came from one study (Shapson and Others, 1978) [27], six comparisons from another (Passarella, 1977) [25], and five deltas from another (Coleman and Others, 1966) [9]. Thus, *one* study produced 42 percent of the total number of comparisons on teacher effect, two studies produced 61 percent of these deltas, and three studies supplied 77 percent of the comparisons. Curiously, this fact also was never addressed in Meta-Analysis II. There can be little doubt that these findings were highly overgeneralized.

INCONSISTENCY IN THE TYPE OF STUDIES USED TO DRAW CONCLUSIONS FROM THE META-ANALYSES

In a November 1976 article in *Educational Researcher*, Gene Glass described the theory and rationale behind the meta-analysis technique of research integration. In that article, Glass wrote at length about the value of using *poorly*-controlled studies, in addition to the best designed studies, in a research data base:

> A common method of integrating several studies with inconsistent findings is to carp on the design or analysis deficiencies of all but a few studies--those remaining frequently being one's own work or that of one's students or friends--and then advance the one or two "acceptable" studies as the truth of the matter. This approach takes design and analysis too seriously, in my opinion. I don't condone a poor job of either; but I also recognize that a study with a half dozen design and analysis flaws may still be valid. Most research criticism I read--and some that I've written--is airy speculation, unbefitting an empirical science. It is an empirical question whether relatively poorly

---

[1/] A count of the individual component variables contained in the data listing that comprised the overall category "teacher effects" comes to 31 comparisons, rather than 30 as reported in the text of Meta-Analysis II.

> designed studies give results significantly at variance with those of the best designed studies; my experience over the past two years with a body of literature on which I will report in a few minutes [psychotherapy outcome research] leads me to wonder whether well-designed and poorly-designed experiments give very different findings. At any rate, I believe the difference to be so small that *to integrate research results by eliminating the "poorly done" studies is to discard a vast amount of important data.* [emphasis added] [13:4]

In the meta-analysis on class size and pupil achievement, Glass and Smith contended that in previous class size reviews "literature searches were haphazard and often overly selective." [16:1] Discussing the findings of Meta-Analysis I in the *Phi Delta Kappan*, the project researchers seemed to sustain the idea presented in Glass' 1976 article when they announced the study as one which used all the vast research literature in coming to its conclusions. In contrast to previous reviews of the class size research, Cahen and Filby said that "Glass and Smith used *all* the available data to develop a continuous distribution of effects and therefore move their analysis beyond the nominal classification of supportive (favoring smaller classes), nonsupportive (favoring larger classes), and inconclusive (failure to reject the null hypothesis)." [emphasis added] [5:495]

In actual fact, Glass and Smith did not use *all* the available research. They used only 110 of 725 comparisons in their final analysis, from only 14 of the 76 studies they collected:

> The nearly 100 comparisons of achievement from the *well-controlled* studies thus form the basis of our conclusion about how class-size is related to academic achievement. [emphasis added] [16:v]

Selectivity using research findings in class size reviews has not ended with the publication of Meta-Analysis I, despite assertions to the contrary. Why did the authors collect 76 studies, making the claim that this large data base "exceeds by 50 percent the most extensive reviews published to date" [16:3], and then use only 14 of them to drawn conclusions from the research? The reason for using only "well-controlled" studies in Meta-Analysis I, according to Glass and Smith, was that there were

> . . . large differences in the class-size and achievement relationship depending on whether pupil assignment was random or uncontrolled . . . . In an area of research where the quality of methodology interacts with the findings of studies, *the results of the best designed studies should be given more weight in drawing conclusions.* The curve for the well-controlled studies . . . , then, is *probably* the best representation of the class-size and achievement relationship. [emphasis added] [16:43]

A logical question then follows: how easy is it to isolate and identify "well-controlled" class size studies? In the meta-analyses, a single criterion was used to define "well-controlled" studies--random assignment of pupils to classes of different sizes. William Murphy, writing for the New England School Development Council (NESDEC), contended that this was not enough:

> The final regression equation and resultant curve generated in this study are based upon 100 comparisons fron ostensibly well controlled studies. Anyone familiar with even a part of the class size literature knows that control of the several independent variables in class size research has been as difficult as in the research on teaching methodologies and effectiveness. Even research claiming randomly assigned or matched groups only controls at best for the student variable. Other independent variables such as teacher competence, instructional method, learning materials used, subject matter taught, and the like often escape control in the best of the studies cited. Well controlled conditions in class group settings are extremely difficult to establish and equally difficult to find in the literature on class size. [24:2]

Unlike Meta-Analysis I, Smith and Glass found more pronounced results in Meta-Analysis II from *"uncontrolled"* studies than from "well-controlled" studies. Unlike Meta-Analysis I, the authors used *all* class size comparisons in reaching their conclusions from the data in Meta-Analysis II. They tried to explain why "uncontrolled" studies had a greater effect on their data than "well-controlled" studies by saying:

> One explanation is that the poorly-designed studies are not credible and that the over-all class-size effect is inflated because 60 percent of the effects come from uncontrolled studies. The more optimistic view is that the effect of class size on the quality of education is a robust effect, detectable even with less sophisticated and powerful research methods. [30:38]

As Glass and Smith repeatedly stated, much of the basis for the conclusions reached in the two meta-analyses hinges on the distinction made by the authors between "well-controlled" and "poorly-controlled" class size studies. However, there is no consistent application of whether or not "well-controlled" studies or "poorly-controlled" studies or all studies should be used in reaching conclusions from the data contained in the meta-analyses. This leads to confusion as is shown when these statements are placed in juxtaposition:

- In his 1976 paper, Glass advocated the use of *all* studies, including "poorly done" research, in the data base of a meta-analysis. To eliminate these "poorly done" studies, said Glass, "is to discard a vast amount of important data."

- In Meta-Analysis I, Glass and Smith based their broad overall generalization on only the *"well-controlled"* studies, saying that "where the quality of methodology interacts with the findings of studies, the results of the *best designed studies* should be given more weight in drawing conclusions." [emphasis added] Yet this rationale seemingly was ignored when in the same meta-analysis *all* studies were used to formulate the conclusion relating to the effect of class size on student grade level.
- In Meta-Analysis II, Smith and Glass based their final conclusions on *all* studies contained in their data base, even though *"uncontrolled"* studies provided more pronounced results on the class size/nonachievement relationship than "well-controlled" studies. The authors gave an indication as to why they thought this phenomenon occurred, but they never explained why their findings were based on *all* studies, when as in Meta-Analysis I, the quality of methodology interacted with the findings of class size studies just as it did in Meta-Analysis II.

INCLUSION OF TUTORING STUDIES AND STUDIES OF VERY SMALL CLASS SIZES IN THE DATA BASE OF META-ANALYSIS I

The inclusion of studies based on tutoring arrangements and very small class sizes in the data base of Meta-Analysis I skewed the results produced in this study, despite the authors' efforts to minimize this problem.

> Concern was expressed by several persons who examined the preliminary analyses that the curve for the well-controlled studies in Figure 4 [see Figure 4 on page 43] might depend excessively on the twenty or thirty comparisons of very small class-sizes (one and two up to five, say) in the data base. When all those comparisons for which S=1 were removed, the curve in Figure 4 for well-controlled studies was even steeper than that shown; this finding is contrary to the claim that tutoring studies skewed the curve unnaturally. When all comparisons for which S was less than 6 were removed, the curve for well-controlled studies became less steep; however, it still rose from the 50th percentile at size 40 to the 60th at size 10, the 67th at size 5 and the 74th at size 1. [16:43, 45]

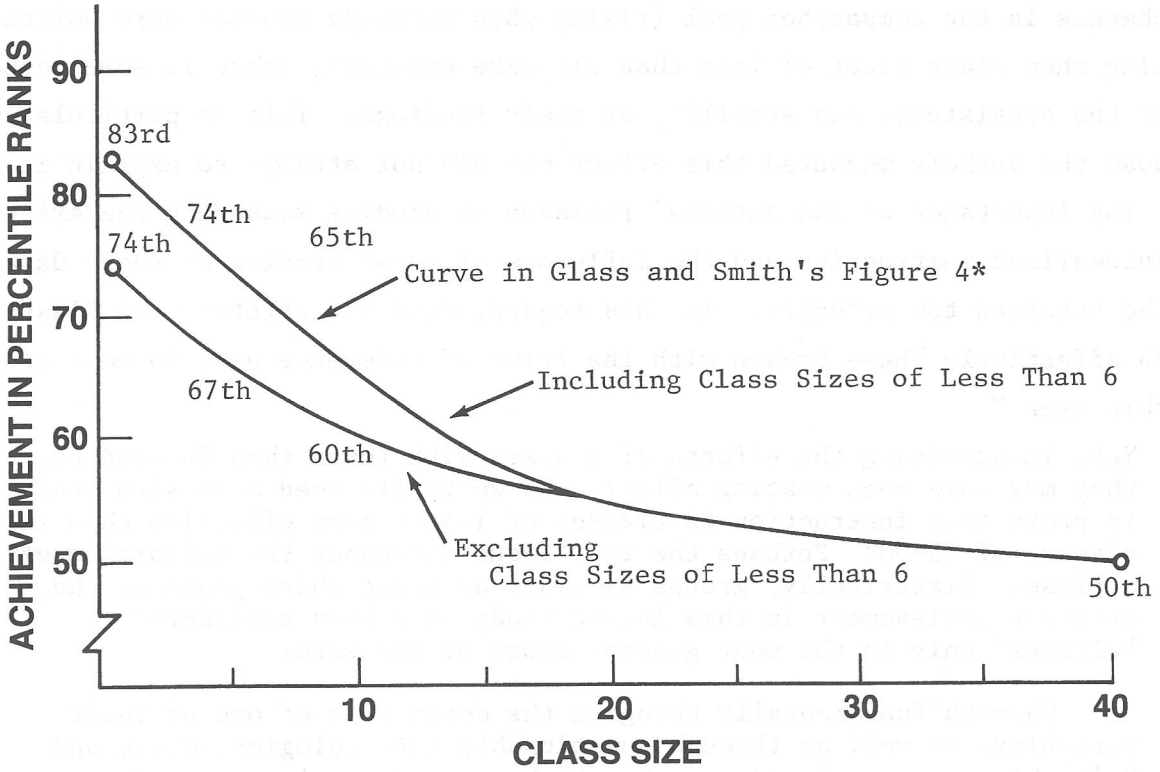
In examining the Glass and Smith statement above on the effect of removing comparisons of S=1 from the data base, William Murphy, writing for NESDEC, observed that:

> As if to minimize the effect of 1-to-1 tutoring on the regression line, the authors state that when comparisons based on a small group of 1 student were omitted, the inverse relationship between size and achievement was even greater! In other words, *the inclusion of the data derived from the smallest of all class sizes tended to diminish the class size effect, and the report takes no special note of this surprising contradiction.* [emphasis added] [24:3]

It is difficult to understand how the authors could recognize, and then dismiss, the effect of including very small class sizes in the data base of Meta-Analysis I when these effects substantially influenced the shape of their final curve. Apparently these effects were so unstable that they did not even vary in the same direction. When class sizes of one were omitted, the curve became steeper; when class sizes of less than six were omitted, the curve became less steep. Nowhere was the cause of this unusual effect or its ramifications discussed.

To imply that the inclusion of very small class size studies in the data base did not influence the shape of the final curve is also to minimize the importance of the percentile rank differences found between the curves in Figure 4. This figure shows the relationship between the curve the authors derived from *including* class sizes of less than six in the data analysis (the final curve based on 14 "well-controlled" studies) and the curve derived from *excluding* class sizes of less than six (plotted from the data points indicated in the statement above quoted from page 45 of Meta-Analysis I). If there was little difference in these curves, as Glass and Smith implied, then the authors also seemed to dismiss the importance of the 9 percentile ranks difference between class sizes of one in these curves, the 7 percentile ranks difference between class sizes of five, and the 5 percentile ranks difference between class sizes of 10. This is especially curious given that one of Glass and Smith's "bold generalizations" was that "the difference in being taught in a class of 20 versus a class of 40 is an advantage of 6 percentile ranks." [16:v] In the second meta-analysis when Glass and Smith quoted from their May 1978 unpublished edition of Meta-Analysis I, they said that "the difference in being taught in a class of 20 versus a class of 40 is an advantage of ten percentile ranks . . . . Few resources at the command of educators will reliably produce effects of that magnitude." [30:2] Why would six or 10 percentile ranks be so important in one situation, but nine percentile ranks not be important at all in another?

FIGURE 4.--The Relationship Between Class Size and Achievement Using Data from "Well-Controlled" Studies: Including Class Sizes of Less Than 6 in the Data Base vs. Excluding Class Sizes of Less Than 6 in the Data Base



*Curve on which the final conclusions of Meta-Analysis I were based (same as Figure 1 on page 6 of this critique).

SOURCE: Gene V Glass and Mary Lee Smith. *Meta-Analysis of Research on the Relationship of Class-Size and Achievement.* San Francisco, California: Far West Laboratory for Educational Research and Development, September 1978. Figure adapted from pp. vi, 45.

Furthermore, adjusting the pool of comparisons to omit tutoring arrangements and classes of less than six pupils affected the curve only for very small classes of less than 20 pupils. Examination of the results published in the study shows that the effects of class size are negligible in the practical mid-range. It appears that the authors' claims regarding the effect of class size on achievement were based primarily on the steepness of the curve for class sizes of less than 20 pupils. Since this region of the curve was most affected by changes in the comparison pool (rising when tutoring studies were omitted and falling when class sizes of less than six were omitted), there is some doubt about the consistency and stability of their findings. This is particularly so, because the authors measured this effect but did not attempt to explain it.

The importance of the authors' reliance on studies measuring the effects of individualized instruction and the influence of these studies on their data cannot be stressed too strongly. In this regard, Hess has written that Glass and Smith effectively "have broken with the frame of reference used in most studies of this type."

> Yet, in examining the effects of classes with fewer than 10 students, they may have been wasting effort. Do we really need a massive study to prove that instruction in classes of 1-5 is more effective than in classes of 25-30? Perhaps the real issue surrounds the definition of a class. Historically, groups as small as those which produced the greatest achievement in this latest study have been considered "classes" only in the most general sense of the word.
>
> Through fundamentally changing the definition of one of their variables, as well as through questionable methodologies, Glass and Smith have thus produced a rather shaky set of conclusions. The bases and techniques of their approach drastically limit the value of its results. [21:7]

POINT 3

Interpretations of Findings Often Contradictory

The findings of the meta-analyses and published interpretations of them in various journals are contradictory on numerous points. These contradictions have been noted specifically in discussions relating to the degree of reductions in class size needed to affect pupil achievement, the effects of class size on achievement in the mid-range, the effects of pupil age and subject matter taught to class size and pupil achievement, and the general effects of class size on the educational environment and on achievement.

In commenting on their findings, Glass and Smith and other members of the Class Size and Instruction Project have contradicted themselves in a number of areas. What appeared in Meta-Analysis I at times bears little if any relationship to what the project researchers have written in summaries of their research in professional journals. Careful reading of these statements leaves one puzzled as to just what conclusions are warranted from these two meta-analyses.

DEGREE OF REDUCTIONS IN CLASS SIZE NEEDED TO AFFECT PUPIL ACHIEVEMENT

The project researchers concluded that little difference in pupil achievement results in class size increases or decreases in the mid-range. Writing in the *Phi Delta Kappan*, project directors Cahen and Filby concluded from Meta-Analysis I that substantial reductions in class size have to be made before pupil achievement is affected:

> The data suggest that there is relatively little pay-off for small overall reductions (e.g., 28 to 25) [in class size]. Attention should be given to ways to make *larger reductions in more limited situations*. [emphasis added] [5:495]

However, three paragraphs later in the same article, they shifted their focus to nonachievement outcomes and made a policy recommendation that is inconsistent with the statement above.

> We need to consider a broad range of outcomes--the relationship between class size and the quality and humanness of the nation's schools. These concerns may make *even small changes in class size worthwhile* and may increase the impetus to find ways to create some small classes. [emphasis added] [5:538]

In the second quote, perhaps Cahen and Filby were taking into account *nonachievement* effects relating to class size. However, Meta-Analysis II, which was concerned with such effects, had not yet been published and was not discussed in the article. Even when Meta-Analysis II was published later, its findings, like Meta-Analysis I, did not provide justification for *small* general decreases in class size typically found in the elementary or secondary grades. A general argument for small changes in class size, whether or not based on achievement or nonachievement effects, is not supported in either the findings of Meta-Analysis I or Meta-Analysis II.

EFFECTS OF CLASS SIZE ON ACHIEVEMENT IN THE MID-RANGE

Another puzzling inconsistency and abrupt shift exists in Glass and Smith's discussions of their findings relating to the effect of class size on the mid-range of pupil achievement. This inconsistency relates to purported class size effects on hypothetical pupil achievement percentile ranks. In Meta-Analysis I, Glass and Smith stated that:

> As class-size increases, achievement decreases. A pupil, who would score at about the 83rd percentile on a national test when taught individually, would score at about the 50th percentile when taught in a class of 40 pupils. The difference in being taught in a class of 20 versus a class of 40 is an advantage of *6 percentile ranks*. [emphasis added] [16:v]

However, this statement should be compared with one made by Smith and Glass in Meta-Analysis II:

> In earlier papers . . . we demonstrated a substantial relationship between class size and achievement. *Those studies which employed rigorous controls* yielded results which taken together, showed that:
>
> > As class-size increases, achievement decreases. A pupil, who would score at about the 63rd percentile on a national test when taught individually, would score at about the 37th percentile in a class of 40 pupils. The difference in being taught in a class of 20 versus a class of 40 is an advantage of *ten percentile ranks*. . . .
> >
> > Glass and Smith, 1978 (p. i) [emphasis added] [30:2]

In Meta-Analysis I, Glass and Smith based their final conclusion on the 110 comparisons from "well-controlled" class size studies. From their analysis, there was an advantage of six percentile ranks between class sizes of 20 and 40. But when Smith and Glass discussed the results of Meta-Analysis I in the introduction to Meta-Analysis II, for some unexplained reason they did not cite their previously *published* finding. Instead, they quoted from page i of their earlier unpublished May 1978 edition of Meta-Analysis I, where their conclusions were based on *all* 76 studies, not on "those studies which employed rigorous controls" as they stated in Meta-Analysis II. In the earlier version of Meta-Analysis I, there were 10 percentile ranks difference between achievement in classes of 20 versus classes of 40. If "the nearly 100 comparisons of achievement from the well-controlled studies . . . form the basis of our conclusion [in Meta-Analysis I] about how class-size is related to academic achievement" [16:v], then why did Smith and Glass quote in Meta-Analysis II from their earlier edition, with conclusions based on all 725 comparisons? This significant inconsistency and unexplained shift in citing their own findings from their own previously published study is most confusing to any careful reader of the two meta-analyses.

EFFECTS OF CLASS SIZE ON ACHIEVEMENT IN THE ELEMENTARY AND SECONDARY GRADES

In Meta-Analysis I, Glass and Smith concluded that "the class-size and achievement relationship seems *consistently* stronger in the secondary grades than in the elementary grades." [emphasis added] [16:40] This statement should be compared with the following three observations, each of which was made later by project researchers in reference to the effects of class size on achievement:

(1) Small-class advantages were *slightly* stronger at the secondary level. It is our opinion that the advantages are too small to lead to a conclusion that elementary pupils would profit less than secondary pupils if class size were reduced. [emphasis added] [5:493]

(2) The Glass-Smith analyses did not find any general interactions in the data; that is, class size effects were *not* noticeably different for children of different ages or abilities or studying different subjects. [emphasis in the original] [5:495]

(3) Small classes may also be more justifiable in the *primary grades*, to get more students off to a good start. [emphasis added] [17:44]

Thus, the conclusion reached in Meta-Analysis I on the effect on class size and achievement, classified by grade level, was completely opposite to its interpretation given by the researchers in the NEA journal *Today's Education*, the source of statement (3) above. Why this is so was never explained. Certainly statements (1) and (2) that appeared in the *Phi Delta Kappan* and statement (3) from *Today's Education* were not based on the conclusions contained in Meta-Analysis I. The statement in *Today's Education* appears more like one of the results of Meta-Analysis II, in which Smith and Glass concluded that class size had a greater impact on the nonachievement effects of elementary-age pupils than secondary-age students. [30:34] However, both of the articles on the first Glass and Smith study published in these two journals appeared well in advance of the publication of Meta-Analysis II which would seem to preclude possible confusion between the findings of these two studies.

Another question arises in analyzing the relationship of class size and achievement on student grade level as presented in Meta-Analysis I. On page 38, Glass and Smith stated that this relationship was calculated for "pupils of age 11 years or younger (elementary) and 12 years or older (secondary)." From the results of regression analysis using 691 comparisons, the authors then found that "the class-size and achievement relationship seems consistently stronger in the secondary grades than in the elementary grades." [16:40] However, when on page 26 they presented a distribution of all comparisons for various ages of pupils included in the data base of Meta-Analysis I, Glass and Smith grouped pupil age according to interval categories that seemingly overlap and, therefore, are incompatible with the categories on which they based their later finding. Among the intervals into which class size comparisons for all studies initially were grouped on page 26 were "9-10," "11-12," and "13-14." From these groupings,

it is unclear how elementary pupils could later be classified as pupils age 11 or younger or secondary students age 12 or older.

EFFECTS OF CLASS SIZE ON ACHIEVEMENT AND NONACHIEVEMENT MEASURES IN DIFFERENT SUBJECT AREAS

Meta-Analysis I found that the class size-achievement relationship was consistently stronger in the secondary grades than in the elementary grades. As shown above, Cahen and Filby interpreted the findings of Meta-Analysis I in the *Phi Delta Kappan* to mean that subject matter made no difference in the effect of class size on achievement. However, a statement in the same article suggests a policy that would ignore both of these conclusions:

> It is interesting to ponder what instruction in schools could be with two professionals teaching 30 pupils, at least for *reading and mathematics* in the *primary grades*. [emphasis added] [5:495]

How the project researchers were able to come to *any* conclusion about the effect of class size on different subjects is hard to discern, given the way their data were classified. In Meta-Analysis I, 47 percent of all comparisons were taken from "all subjects combined (i.e., elementary school classes)." [16:24] In Meta-Analysis II, 57 percent of all comparisons were derived from "all subjects," which the authors said came from "experiments with all-day self-contained classrooms." [30:11, 12] Very few class size researchers have ever presented the findings of their research in terms of an aggregate subject-area category such as "all subjects combined," no matter what grade level was examined. Homogenization of data in this way seems to preclude the two meta-analyses from drawing any specific distinctions at all relating to class size effects on subjects used in the existing research. Typically in class size studies pupils are taught and tested in specific subject areas such as reading, mathematics, or English. How or why these subject areas were combined through meta-analysis, and the effects smoothed out, is left unexplained.

GENERAL EFFECTS OF CLASS SIZE ON THE EDUCATIONAL ENVIRONMENT AND ON ACHIEVEMENT

In reading the broad conclusions reached in Meta-Analysis II, one wonders whether or not the conclusions reached were "bold generalizations" or rather "timid qualifications":

> (1) These findings indicate that there is a beneficial effect on the general quality of the educational environment resulting from decreasing class size. [30:27]
>
> (2) Using meta-analysis permits us to unravel the complexity [of the research] and reveal the *small* but consistent effects of class-size. [emphasis added] [30:39, 45]

If the authors classified the effects of class size on nonachievement measures as "small," then one also could say that the effects of class size on achievement were "very small," since achievement effects had less interaction with class size than nonachievement effects. *Even if the results of the meta-analyses were accepted as stated, class size would have only a very small impact on pupil achievement.* However, it is puzzling as to why Smith and Glass, in the opening pages of Meta-Analysis II, said that in the meta-analysis on class size and achievement "we demonstrated a *substantial* relationship between class size and achievement. As class size increases, achievement decreases." [emphasis added] [30:ii]

POINT 4
Confuses Class Size Issue and Unjustifiably Encourages General Class Size Reductions

> The conclusions presented in the meta-analyses only confuse the class size issue and fail to provide practical guidelines or help for making class size decisions. Some of the policy suggestions advocated in the meta-analyses and in subsequent discussions by the project's staff members are unsupported by their own studies findings. Furthermore, proponents of general class size reductions have been unjustifiably encouraged to use the findings of these meta-analyses to support their positions.

To repeat the basic thrust of this critique, Meta-Analyses I and II do not provide convincing evidence to support the "bold generalizations" made that smaller classes are better, under any and all circumstances, in terms of either pupil achievement or the quality of education. It is now important to examine how the project researchers stated the policy implications of their studies.

In Meta-Analysis II, Leonard Cahen, the principal investigator of the Class Size and Instruction Project, commented that:

> The two reports [Meta-Analyses I and II] confront educational decision-makers with reasonable evidence that reduced class size can have positive effects upon classroom processes and pupil learning. *If this evidence is convincing, educators must find ways to reduce class size for at least parts of the school day and year.* [emphasis added] [30:i]

From the criticisms of the use of the methodology and the numberous contradictory statements contained in the meta-analyses and in subsequent published discussions relating to them, this evidence is *not* convincing, especially in relation to the impact of class size on pupil achievement. Moreover, the research evidence to support the argument for smaller classes "for at least parts of the

school day and year" is nonexistent. Which classes should be reduced? "For reading and mathematics in the primary grades," as Cahen suggested in the *Phi Delta Kappan*? [5:495] From data reported in the meta-analyses, it is impossible to say, since nearly half of all comparisons in Meta-Analysis I and almost 60 percent of all the deltas in Meta-Analysis II were taken from "all subjects combined." [16:24; 30:12] Also, Meta-Analysis I concluded that class size had a greater impact on the achievement of *secondary* students [16:40], but Meta-Analysis II found that class size had a greater effect on the "quality of education" of *elementary* pupils. [30:34] The essential question remains: Where is the evidence for such statements? These policy recommendations based on the two meta-analyses are unsupported by the results of these studies.

The policy implications attributed to the findings of Meta-Analysis I have been so exaggerated as to make them less than credible. As a case in point, the October 29, 1979 issue of *Education Daily* contained an article in which Nikola Filby, co-director of the Class Size and Instruction Project, described the results of Meta-Analysis I. In discussing the policy implications of these findings, Filby remarked that: "Theoretically, the money spent for smaller classes would be worth the reduction in crime and welfare costs." [29] Albert Shanker, president of the New York City teachers union and the American Federation of Teachers, also used the two meta-analyses to argue for smaller classes. If average class size is increased from 29 to 31 in New York City as proposed by the mayor to help cut costs, Shanker warned that the consequences include "more middle class taxpayers moving out, businesses relocating to places where they can get educated workers, higher costs for welfare and for crime protection--and, of course, there will be the human misery of the uneducated." [26:E9]

William Murphy pointed out that a collective analysis of many class size comparisons is essential to find a trend favoring small or large classes when meta-analysis is used. Yet Glass and Smith found only a 55:45 chance of discovering a result favoring small classes in the class size literature. Does it then follow that the effects of class size in practical school situations would be just as elusive? Murphy then added:

> That is, would superior achievement in smaller classes in schools be detectable only through a meta-analysis of the pooled comparisons among classes in the school or school system? Would practical advantages of smaller class size have only a 55:45 chance of appearing in a typical school? Such questions warrant answers before expectations for small classes out-distance the actual student achievement. [24:2]

Closely associated with the class size issue are the problems of financing smaller classes. Many class size reviewers have addressed this concern. For example, in its 1978 summary of research on class size, ERS concluded that "even small system-wide changes of one or two pupils per class can have major impact on a school system's budget." [6:70] That the project researchers also acknowledged this fact was illustrated in a quotation from the *New York Times:*

> Mr. Cahen said "it would be economic suicide" to recommend cutting every classroom's enrollment in half, particularly in a budget-cutting atmosphere. The experimenters nevertheless believe that arrangements within a school can be made more flexible, with pupils attending small classes for part of the day. [19]

It appears that the only policy alternative left to the project researchers was to advocate smaller classes for part of the school day and year, an implication that was not even remotely addressed in either meta-analysis. In proposing this policy strategy, other factors also become important. Discussing the effects of arranging small weekly group sessions in order to provide lower pupil-teacher ratios, Fritz Hess noted that additional intervening variables such as the length of time between small classes and the interactive effect of small and large classes within the same subject would be introduced. "The influences of these and other factors," Hess stated, "have not been investigated within the Glass-Smith analysis." [21:6-7]

Moreover, Cahen and Filby expressed dismay over what readers could justifiably infer from the findings of the two meta-analyses:

> We are concerned that the Smith-Glass curve may be interpreted by "budget at any cost" school administrators and citizens to mean that class size can be increased beyond 30 pupils without achievement deficit or other consequences. [6:495]

Robert McClure, a program manager with the National Education Association's Instructional and Professional Development Division, expressed similar concerns in *Today's Education* relating to the implications for class size policy in the mid-range as presented in Meta-Analysis I. It should also be noted that his statement appeared *before* the results of Meta-Analysis II, which dealt with non-achievement measures, were published.

> [Glass and Smith] find only small differences in achievement between classes of 20 and 40 pupils, while classes that contain fewer than 20 students differ greatly from larger classes. It would be exceedingly unfortunate, however, if school boards were to use these data to argue, "Since we cannot afford to get classes below 20, they might as well be at 40." This would be a simplistic answer to a complex question that affects, through the schools, the

> whole society. Such public leaders, to act responsibly, should help their communities understand that quality of schooling is measured in many ways. Small classes produce superior results on all dimensions. Furthermore, as all teachers know, the small differences in achievement may be very important to the pupils involved. [17:43]

Another important point that needs to be stressed is that "monetary factors do not play a major role in the research design of Glass and Smith," as Hess noted in his critique of Meta-Analysis I:

> A review of the implications of the study demonstrates how small this role really is. The authors boldly conclude that a direct relationship exists between class size and achievement, and that the achievement curve begins to rise rapidly only when class size reaches 15 students. Even if one assumes the methodology of the study to be valid, however, financial considerations make the teaching of classes with 15 students difficult at best. Given the current state of educational finance, it is often a challenge to budget for classes of 20-25 students. If the authors' advice is followed one step further, and classes are reduced to between 1 and 19 students, the levels of funding required increases geometrically. [21:6]

Given the extensive problems with the Glass and Smith studies described in this critique, it is disturbing to see how the results of the meta-analyses are being used to influence educational policy. Many apparently are reading the reported findings of the meta-analyses uncritically and therefore believe that the conclusions reached are supported by the meta-analyses. This is understandable, but it presents serious implications. For example, in their November 1979 *Phi Delta Kappan* article that summarized the last decade of educational research, Walberg, Schiller, and Haertel seem to accept the findings of Meta-Analysis I without qualification:

> Gene Glass and Mary Smith's very extensive analyses, moreover, reveal that studies that randomly assign students to small and large classes in true experiments show stronger positive benefits for smaller classes. This finding enhances confidence that smaller classes lead to greater achievement rather than that both are caused by other variables such as community wealth. Stronger size/learning relationships found by Glass and Smith in studies carried out after 1960 than in those before 1940 indicate the increasing sophistication of educational research. Although the inverse size/learning relationship is not the strongest or most consistent among the results summarized here, several estimates from the Glass and Smith work are impressive: Children who gain 1.0 grade equivalents on average per year in a class of 40 would gain 1.3 equivalents in a class of 20 and 1.6 if taught individually. If average pupils were taught in a class of 20 pupils from kindergarten through grade 6, they would be over two years ahead of similar pupils taught for the same length of time in a class of 40. [31:180]

It is impossible to justify from either meta-analysis the statement that "children who gain 1.0 grade equivalents on average per year in a class of 40 would gain 1.3 equivalents in a class of 20 and 1.6 if taught individually." This assumes a precision that is lacking. In the first place, the proper use of the concept of grade equivalents precludes such a sweeping generalization. Grade equivalents are measures of pupil progress calculated on the basis of the range of standardized scores obtained by administering a specific test designed for a specific subject or skill area, to specific groups of students, and at similar grade levels. One of the problems in achievement testing is how to relate properly pupil scores on one standarized test to those of another standardized test covering the *same* subject or skill area. But nowhere in Meta-Analysis I was "achievement in percentile ranks" related to any specific subject or skill area (no distinction was found for subject taught) and nowhere were such hypothetical metrics related to the range of scores on any specific standardized test.

Moreover, there is no support in either of the two meta-analyses for statements such as "if average pupils were taught in a class of 20 pupils from kindergarten through grade 6, they would be over two years ahead of similar pupils taught for the same length of time in a class of 40." Nevertheless, such sweeping conclusions by responsible persons have been made and are encouraged by the overgeneralizations and confusion caused by the published findings of the meta-analyses.

Likewise, it is interesting to note the divergence of opinion between Glass and Smith and Walberg and his associates when they addressed in separate reports the implication of the class size effects on achievement found in Meta-Analysis I. In a quote from their May 1978 edition of Meta-Analysis I, cited by Smith and Glass in the second meta-analysis, that found 10 percentile ranks difference in achievement for pupils taught in a class of 20 versus a class of 40, the authors proclaimed that: "Few resources at the command of educators will *reliably* produce effects of that magnitude." [emphasis added] [30:2] However, acknowledging that "several estimates from the Glass and Smith work are impressive . . . ," Walberg, Schiller, and Haertel also said in the previously quoted statement that "the inverse size/learning relationship is not the strongest or most consistent among the results summarized here." [31:180]

The political implications surrounding class size policy continue to be one of the most significant of all the factors involved in this controversial issue. Many class size reviewers have noted this fact, including Glass and Smith:

> Because the research evidence appeared conflicting, the debate over increasing or decreasing class size has become more political than scientific. Constituencies pull one way or the other, each marshalling that part of the evidence that supports its own case. The decisions eventually made on class size are determined less by evidence than by which side has the greater political power. [30:2]

The weekly newsletter *Education U.S.A.* also acknowledged this fact and cited ways that the results of these meta-analyses presently are being used.

> One point is clear--the size of classes is leaping from a research to a political issue. The NEA and its state affiliates trumpeted the first Glass-Smith report as "conclusive" evidence that schools should hire more teachers. NEA's Bernard McKenna says the Glass-Smith studies "show definitely" that small classes make a difference: "It's what teachers have known and said for a long time." [1:16]

An article published by the Pennsylvania State Education Association (an NEA affiliate) illustrates how some teacher groups are rallying around these findings to support their campaign for smaller classes:

> Despite the hollow claims of school board members interested in being reelected and administrators afraid to admit otherwise, the smaller the class the better the chance that a student in that class will learn more from his or her teacher . . . .
>
> Of course, school boards and chief school administrators always have fallen back on their favorite ploy when their backs are to the wall. They shout to teachers, "Prove it with research." As if the sun rises and falls on data, statistics, and numbers.
>
> Until recently, the trick worked. All teachers and students could say was, "I don't have data or hard facts, but I know that I can teach (or learn) more when I have more time with my students (teacher) on a one-to-one basis."
>
> School boards in the past would sneer and scoff at such statements because they lacked so-called "hard, usable, touchable, computer-like DATA."
>
> But not anymore. The following NEA News Service report calls attention to a recent survey [sic] which now shows--in the form school boards have always demanded--that smaller class sizes DO MAKE A DIFFERENCE! [emphasis in the original] [2]

Further examples illustrate the political impact of the Glass and Smith studies. In a February 1979 budget session held by the school board of Montgomery County, Maryland, a district with an enrollment of over 100,000 students, class size became one of the focal points of the discussion. The *Washington Post* reported that a school board member quoted from the Glass and Smith meta-analysis on class size and pupil achievement in support of limiting the size of classes:

"A clear and strong relationship between class size and achievement has emerged. There is little doubt that . . . more is learned in smaller classes." [20:B-2] On the basis of the school board's action, according to the *Post*, it was agreed to:

- Add 72 teachers throughout the system.
- Drop 38 assistant principals, teachers' aides, and secretaries.
- Limit class size to 28 pupils in grades 1-3, 30 in grades 4-6 and 32 in secondary school classes. [20:B-2]

In a paid advertisement that appeared in the *New York Times* on February 23, 1980, Albert Shanker argued that recent small increases in average class size proposed by New York Mayor Koch were unjustified, based on the class size research.

> When Koch announced huge cuts in the school budget, he said that (1) his budget cuts would result in small increases in class size, with the average class rising from 29 to 31, and (2) research shows that small class size isn't important to good education.
>
> Koch is wrong on both counts. . . . What evidence is there that he is wrong?
>
> First, there is plain common sense. . . .
>
> But common sense is supported by substantial scientific research. [26:E9]

Shanker then said that Meta-Analysis I "reviewed 80 different studies and concluded that: 'As class size increases, achievement decreases . . . . The difference in being taught in a class of 20 versus a class of 40 is an advantage of 10 percentile ranks . . . . Few resources at the command of educators will reliably produce effects of that magnitude.'" He also quoted extensively from Meta-Analysis II, to the effect that class size affects the quality of the classroom environment, pupils' attitudes, and teachers. As discussed previously in this critique, the result of increasing average class size from 29 to 31 in New York City, in Shanker's words, "will mean more middle class taxpayers moving out, businesses relocating to places where they can get educated workers, higher costs for welfare and for crime protection--and, of course, there will be the human misery of the uneducated." [26:E9]

These are examples of some of the things that have happened when the Glass and Smith findings have been used in an effort to make the case for small classes. Policy makers would be misled if they were to rely on the authors' statement that "few resources at the command of educators will reliably produce effects of that magnitude." [30:2] This critique has shown repeatedly that, despite appearances,

the meta-analysis conclusions are neither "substantial" nor "reliable." These examples show the potential risk of uncritical acceptance of these meta-analyses.

In a review published by the New England School Development Council that critiqued the Glass and Smith meta-analysis on class size and achievement, William Murphy argued that: "Given the realities of public education, the press for austerity and for accountability, the Glass-Smith study offers little more to guide policy decisions in the area of class size than has the empirical research on which it is based." [24:5] It should be clear that the use of class size research findings still calls for prudence, rather than zeal, on the part of school policy makers. The conclusions reached in the ERS Research Brief on class size stand as valid today as they did before the publication of the Glass and Smith meta-analyses, including these ERS findings:

- In terms of pupil benefits, research findings fail to justify small overall reductions in class size or pupil-teacher ratio by a school board merely as a matter of general policy *without definite pupil-benefit objectives for specific groups of pupils.* [emphasis added] [6:70]

- Policy decisions relating to class size and pupil-teacher ratio involve factors that are complex, varied, and often emotionally charged. These require the weighing of the possible pupil benefits, the possible teacher benefits, the facilities utilized, the financial costs, and the possible political consequences. [6:70]

POINT 5
Creates Doubt About the Need for Further Research

> By overgeneralizing their data and drawing "bold generalizations" where only cautious and qualified conclusions were supportable and by implying that they have put the class size issue to rest, the authors have created doubt about the need for conducting further research into the class size issue.

For nearly a century, educational researchers have investigated the class size question. Until the publication of the two Glass and Smith meta-analyses, there was no "last word" on the subject. Researchers who examined the class size literature over the years have found it nearly impossible to take a solid stand in favor of either smaller or larger classes as a matter of general policy and have usually called for more research. However, since the appearance of the two meta-analyses on class size, some persons and groups have accepted these findings as the ultimate in class size research, apparently assuming that further research into this highly complex and controversial issue is no longer needed. For example, Robert McClure of NEA's Instruction and Professional Development group was quoted in *Today's Education* as saying:

> The research reported in this landmark study [Meta-Analysis I] is of great significance to teachers. Many policymakers and other researchers, using less sophisticated methodologies, are claiming that class size is relatively unimportant to pupil growth. *That argument can now be put to rest.* This research demonstrates that smaller classes do have a positive relationship to student achievement. [emphasis added] [17:43]

In Meta-Analysis I, Glass and Smith themselves implied that their analysis was the ultimate in the long history of class size research:

> In the research reported here, an attempt was made to correct these shortcomings [that were present in earlier class size studies, according to Glass and Smith] and determine if the huge research literature on class-size and achievement really was hopelessly confusing or if its message was merely buried in myriad results waiting to be coaxed out with more advanced methods of research integration. [16:2]

Yet given the many criticisms of Meta-Analysis I and Meta-Analysis II and the interpretations of these results that have been examined in this critique, the two Glass and Smith meta-analyses are far from being the last word on the subject. In fact, the authors have missed the central point of the entire class size issue by addressing the question as simply: Are smaller classes better? They then answered with an emphatic and unqualified: YES! To address the class size issue in such simplistic terms and then answer in such a simplistic manner can lead to costly class size decisions that not only fail to achieve their intended purpose of improving instructional programs but also can result in disservice and even harm to those pupils who might otherwise benefit the most from smaller classes.

The real question in the class size issue is: What types of students might benefit the most from smaller classes and under what conditions? Although much more research is needed before this question can be answered definitively, in its summary of research on class size, ERS attempted to draw *tentative* conclusions from the available research in an effort to provide the best guide for class size decisions possible at this time. These tentative conclusions point out that smaller classes *can* have positive effects on pupil achievement in reading and mathematics in the early primary grades and for low-ability, economically disadvantaged, and socially disadvantaged pupils. Conversely, larger classes may be justified in areas in which pupil achievement is affected little and where facilities and personnel may be used more efficiently. The reality of limited resources requires that special attention to special needs and the use of effective means of meeting them must be targeted toward specific pupils in specific subject areas and in specific grade levels. A blanket approach to class size policy can be not only expensive but also unjustified so far as existing research would indicate. The question of efficient class size should not be viewed as final and absolute, but as emerging and varied.

The results of these two reports, presented as unarguable fact, proclaimed the idea that smaller classes in themselves are more beneficial than larger classes, a finding not made in such absolute terms in other reviews of the class

size research. Yet as this critique has shown, such a "bold generalization" is unjustified from either meta-analysis. Moreover, school officials and others concerned with pressing class size decisions can become frustrated and feel they are without guidance from the research. The end result of these unwarranted generalizations may well be to cast even dimmer light on the value of further research to improve schools and educational programs for pupils and the need to support such research.

It would be unfortunate if this proved to be the case, for much research is needed on the effects of class size on both pupil achievement and nonachievement processes. Further research is needed that would focus first on those groups of pupils and subject areas already tentatively identified in the existing class size research as being positively affected by smaller classes: pupils in the primary grades, economically disadvantaged pupils, socially disadvantaged pupils, low ability pupils, and in the areas of reading and mathematics. Research should be designed so that experimenters and practitioners will be able to learn more about the efficient ranges of class sizes and conditions of learning for such pupils, especially in these subjects.

There is also need for research concerning efficient class sizes and teaching techniques in other subject and skill areas. The effective allocation of resources among pupils, teachers, classrooms, subject areas, and grade levels pose major problems that deserve much attention from research. There are other class size questions that need research with improved research technology. *But it is most important that the research should be as specific and as practical as possible.*

Class size is not and should not be a dead issue for educational researchers, despite statements or implications to the contrary. To imply that the two meta-analyses on class size represent the final word on the issue does a disservice both to education and to research.

SUMMARY AND CONCLUSIONS

The two meta-analyses of class size research by Gene V Glass and Mary Lee Smith and published by the Far West Laboratory for Educational Research and Development have been widely interpreted as providing final and convincing evidence that smaller classes are better than larger classes. In their first report, Glass and Smith stated emphatically that their findings contradicted the conclusions reached by previous analysts of class size research.

Because of the use of what they termed "sophisticated" meta-analysis (described as an "analysis of analyses") and the inclusion of a large volume of class size studies, Glass and Smith claimed that their findings were superior to all previous class size reviews and that they were able to make "bold generalizations" where previous analysts were able to make only "timid qualifications." They stated unequivocally that their first meta-analysis "established clearly that reduced class-size can be expected to produce increased academic achievement." [16:iv]

In the process of examining the two meta-analyses, ERS was surprised and concerned to find that several of the claims and conclusions made in these reports and in interpretations of them were unsupported or unjustified. A number of the interpretations of the findings and the recommendations for educational policy were not only unsupported but also conflicting.

According to Glass and Smith, the basic finding of their meta-analysis on the effect of class size on pupil achievement was shown in a graph that has been widely published and presented as conclusive evidence that smaller classes result in increased achievement. However, the graph shows that in the wide range of 20 to 40 pupils, class size makes little difference in achievement. The authors stated that "the major benefits from reduced class-size are obtained as size is

reduced below 20 pupils." [16:v] It is important to note that a substantial proportion of the comparisons used to influence the part of the graph below 20 pupils dealt with extremely small instructional arrangements, such as one-to-one tutorial arrangements and classes of 2 to 5 pupils. Furthermore, the authors stated that the class size and pupil achievement effect was not influenced by "'source of data,' 'subject taught,' 'duration of instruction,' 'pupil IQ,' and 'type of achievement measure.'" [16:38]

The methodology used in these meta-analyses purportedly was designed to correct the imprecise reliance on the findings of a few studies by previous reviewers of class size research. Actually, the major findings of Meta-Analysis I, which examined the effects of class size on pupil achievement, rest on *fewer* studies, not more. The central findings of Meta-Analysis I were not based on all 76 studies analyzed by Glass and Smith but only on 14 studies that the authors considered experimentally "well-controlled" solely because pupils were assigned randomly to classes. Of these 14 studies, only 6 pertained to class size situations that typically apply in elementary or secondary schools. In Meta-Analysis II, which examined the effects of class size on nonachievement processes, only 6 studies provided 43 percent of the total comparisons.

On the surface the two meta-analyses appear to be statistically sophisticated and to provide new insights; actually they are far too insensitive to detect many of the important relationships relating to class size. The methodology used obliterated significant distinctions contained in the research; thus important findings were lost in the process. When carefully and cautiously analyzed, the available research on class size provides many more indicators for decisions concerning the effective and appropriate grouping of pupils and teachers for instructional purposes than the two Glass and Smith meta-analyses reveal.

In the meta-analyses, serious problems were created by attempting to homogenize diverse statistical data; this technique makes it necessary for the reader to be placed in what Glass and Smith termed the "benign jeopardy of losing qualitative and personal familiarity with the research." [16:4] But by preventing readers from becoming familiar with the class size research being analyzed, the authors place them in the "benign jeopardy" of being unable to examine or evaluate the conclusions reached through meta-analysis. Persons concerned about class size decisions need to become *more* familiar with the results of class size research, not less.

Numerous contradictions were found in the findings of the published meta-analyses and in their interpretations in various journals made by the Class Size and Instruction Project researchers. These contradictions were noted in discussions pertaining to the nature of class size reductions needed to affect pupil achievement, the effects of class size on achievement in the mid-range, the effects of pupil age and subject matter to class size and achievement and to non-achievement measures, and the general effects of class size on the educational environment.

The conclusions from these meta-analyses only confuse the class size issue; at the same time, they fail to provide practical guidelines or help for making class size decisions. Some of the policy suggestions advocated from the use of these conclusions are without support from the meta-analysis reports. Furthermore, proponents of general class size reductions are unjustifiably encouraged by the conclusions of the two meta-analyses.

By unequivocally announcing "bold generalizations" and by implying that they have put the class size issue to rest, the authors have created doubt about the need for conducting further research into the class size issue.

"Bold generalizations" are of little value in deciding specific class size policy. They tend to result in rigid and stereotyped action. What is needed are practical guidelines for flexible class size policy. Flexibility allows decision makers to vary the size of classes to fit the needs of pupils, teachers, and diverse school situations. It can be damaging to assume, as do these meta-analyses and the interpretations of them, that class size policy should be determined by the simple generalization that smaller classes are always better. School officials, educators, and others would be better advised to go beyond such generalizations regarding the effects of class size and to address the question: What types of students might benefit the most from smaller classes? Moreover, the sweeping generalizations made in these meta-analyses give no direction for further research that might be helpful in making class size decisions relating to the special needs of pupils.

A number of previous reviews of class size research using traditional approaches have provided more helpful and practical guides for use in making class size decisions than have the two recent meta-analyses. In addition, these previous reviews also have provided indicators that can help focus future class size research on important issues.

From Meta-Analysis I Glass and Smith could find a difference of only 6 percentile ranks between the average achievement of pupils in classes of size 20 and those of size 40. ERS sees nothing in this finding, or elsewhere in the meta-analysis, to contradict the conclusion contained in the 1978 ERS Research Brief on class size that "within the mid-range of about 25 to 34 pupils, class size seems to have little if any decisive impact on the academic achievement of most pupils in most subjects above the primary grades." [6:69]

Similarly, ERS finds nothing in Meta-Analysis I to change its conclusion that smaller classes *can* have a positive influence on pupil achievement in reading and mathematics in the early primary grades and for low-achieving and economically or socially disadvantaged students.

Further research is needed that would focus on such specific groups of pupils and subject areas which have been tentatively identified as being positively affected by smaller classes. It is most important, however, that such research be as specific and practical as possible.

Moreover, ERS finds nothing in Meta-Analysis II to change its conclusion that few pupil benefits can be expected from reducing class size if teachers continue to use the same teaching methods in smaller classes that they used in larger classes. Nor does ERS find evidence to change its conclusion that when classes are reduced, many teachers do not take advantage of them to individualize instruction. Thus, there are many class size questions that need further research with improved research technology, such as those relating to class size and teaching procedures, teacher motivation, and cost/quality relationships.

In the final analysis, ERS finds that, despite claims to the contrary, the two meta-analyses fail to provide any new evidence relating to class size research that holds important implications for educational policy. Furthermore, it would be a mistake for educational policy makers to rely on the Glass and Smith conclusions when deciding class size policy.

It should be emphasized that the purpose of the critique has not been to make a case for either smaller or larger classes, but rather to analyze the findings and interpretations of data from the two meta-analyses of class size research that have been published recently. By so doing, ERS hopes to continue to help school officials, educators, and others to interpret class size research and to make sound decisions pertaining to class size.

APPENDICES

APPENDIX A

Bibliography of the 14 "Well-Controlled" Studies Used for the Final Conclusions of Meta-Analysis I

Study ID Number		Number of Comparisons
16	Balow, I.H. "A Longitudinal Evaluation of Reading Achievement in Small Classes," *Elementary English*, 46 (February 1969), pp. 184-187. (Journal listed incorrectly in Meta-Analysis I as *"Elementary Education."*)	2
1	Bausell, R. B., W. B. Moody, and F. N. Walzl. "A Factorial Study of Tutoring Versus Classroom Instruction," *American Educational Research Journal*, 9 (1972), pp. 591-598.	3
73	Bostrom, E. A. "The Effect of Class Size on Critical Thinking Skills." Ed.D. thesis, Arizona State University, 1969. (UMI Order No. 69-1276)	1
52[1/]	Cram, B. M. "An Investigation of the Influence of Class Size Upon Academic Attainment and Student Satisfaction." Ed.D. thesis, Arizona State University, 1968. (UMI Order No. 68-14988) Jeffs, G. A. and B. M. Cram. *The Influence of Class Size on Academic Attainment and Student Satisfaction.* Las Vegas, Nevada: Ed W. Clark High School, 1968. (ED 021 252)	3
8	DeCecco, J. P. "Class Size and Co-ordinated Instruction," *British Journal of Educational Psychology*, 34 (February 1964), pp. 65-74.	2
61	Ellson, D. G. and Others. "Programed Tutoring: A Teaching Aid and a Research Tool," *Reading Research Quarterly*, 1 (1965), pp. 77-127.	2
6	Haskell, S. "Some Observations on the Effects of Class Size Upon Pupil Achievement in Geometrical Drawing," *Journal of Educational Research*, 58 (1964), pp. 27-30.	4

Study ID Number		Number of Comparisons
49	Meredith, V. H., L. M. Johnson, and R. A. Garcia-Quintana. *South Carolina First Grade Pilot Project 1976-77: The Effects of Class Size on Reading and Mathematics Achievement.* Columbia, South Carolina: South Carolina Department of Education, January 1978.	4
9	Moody, W. B. and Others. "The Effect of Class Size on the Learning of Mathematics: A Parametric Study," *Journal of Research in Mathematics*, 4 (1973), pp. 170-176. (ED 062 138)	6
55	Ronshausen, N. "The Programed Math Tutorial--Paraprofessionals Provide One-to-One Instruction in Primary School Mathematics." Paper presented at the annual meeting of the American Educational Research Association, 1975. (ED 106 743)	8
3	Shaver, J. P. and D. Nuhn. "The Effectiveness of Tutoring Underachievers in Reading and Writing," *Journal of Educational Research*, 65 (1971), pp. 107-112.	18
77	Smith, D. I. "Effects of Class Size and Individualized Instruction on the Writing of High School Juniors." Ph.D. thesis, Florida State University, 1974. (UMI Order No. 74-25461)	18
58	Verducci, F. "Effects of Class Size Upon the Learning of a Motor Skill," *Research Quarterly*, 40 (May 1969), pp. 391-395.	3
35	Wright, E. N. and Others. *Effects of Class Size in the Junior Grades.* Toronto, Ontario: Ontario Ministry of Education, 1977.	36
14 =	total number of "well-controlled" studies	total number of comparisons in "well-controlled" studies = 110

1/ Although Glass and Smith stated that 77 studies comprised the total data base of Meta-Analysis I, the bibliography and the data listing actually contain 76 separate studies. Two of the entries (Cram [1968] and Jeffs and Cram [1968]) reported data dealing with the same class size experiment.

In their May 1978 edition of Meta-Analysis I, Glass and Smith assigned both of these reports the same Study ID Number (52), but counted each study separately in arriving at the total number of 77 class size studies they examined. Because these reports presented the same data, ERS counted them only once in the computation of the number of total studies and the number of "well-controlled" studies analyzed by Glass and Smith. Thus, there are 76 studies in the total data base (although there are 77 separate entries) and there are 14 "well-controlled" studies (although there are 15 separate entries).

APPENDIX B

Summary and Conclusions from the 1978 ERS Research Brief on Class Size

The following is the full text of the "Summary and Conclusions" section contained in the ERS Research Brief, *Class Size: A Summary of Research* (1978), pp. 68-70. It is reprinted here to provide reference to the tentative conclusions reached by ERS in its 1978 analysis of class size research.

The continuing debate over class size poses hard questions for school policy makers and school administrators today. Most reviews of the research have found the overall effects of class size on pupil achievement to be inconclusive--some studies reported that smaller classes were better, some that larger classes were more effective, while others could find no difference between the two. For more than two decades, the broad question "Do smaller classes result in increased educational quality?" has dominated the debate concerning class size. Frequently the number of studies favoring smaller classes have been counted and compared to the number of studies favoring larger classes, with many costly decisions having been based on the outcome.

This analysis of original research studies and reviews of research on class size and pupil-teacher ratio indicates clearly that the cause and effect relationships pertaining to the class size issue are highly complex and interlocked with many other variables.

The research provides no clearcut guidelines for an "optimum" class size covering all types of students at all grade levels. Students at different levels of personal and academic development require different learning conditions in order for optimum gains in achievement to occur. Therefore, it may be more beneficial for school officials, educators, and others to go beyond the customary generalities regarding class size and address the question of: "Which types of students might benefit the most from smaller classes?"

It is difficult to summarize the results of the array of studies and reviews of research on class size included in this Research Brief. Certainly more careful research is needed to answer conclusively many of the important questions about the effects of class size on pupil learning, on the teaching process, on teacher morale and job satisfaction, and on cost/quality relationships. But more specific data are not available and yet many important decisions regarding class size cannot wait. Thus, we must carefully draw conclusions from the data at hand. The accumulated evidence to date would appear to support the following *tentative* conclusions for consideration when school officials formulate educational policy:

- Research findings on class size to this point document repeatedly that the relationship between pupil achievement and class size is highly complex.

- There is general consensus that the research findings on the effects of class size on pupil achievement across all grade levels are contradictory and inconclusive.

- Research to date provides no support for the concept of an "optimum" class size in isolation of other factors. Rather the indicators are that efficient class sizes are a product of many variables including: subject area, nature and number of pupils in the classroom, nature of learning objectives, availability of materials and facilities, instructional methods and procedures used, skills and temperament of the teacher and support staff, and budgetary constraints.

- Existing research findings do not support the contention that smaller classes will of themselves result in greater academic achievement gains for pupils. The evidence is that within the mid-range of about 25 to 34 pupils, class size seems to have little if any decisive impact on the academic achievement of most pupils in most subjects above the primary grades.

- There is research evidence that small classes are important to increased pupil achievement in reading and mathematics in the early primary grades.

- There is also some evidence of a positive relationship between small class size and pupil achievement when primary grade pupils are taught in small classes for two or more consecutive years.

- There is evidence that pupils with lower academic ability tend to benefit more from smaller classes than do pupils with average ability.

- Some research indicates that smaller classes can positively affect the scholastic achievement of economically or socially disadvantaged pupils.

- Research on class size suggests the importance of an emphasis on the methods and quality of instruction in the classroom rather than on the quantity of pupils in the classroom.

- There is considerable and consistent research evidence that certain teaching procedures and practices perceived by some educators as conducive to a productive learning environment (e.g., more individualization, creativity, group activity, and interpersonal regard) occur more frequently in smaller classes than in larger classes. But not enough research has been done to validate the presumed superiority of these activities in terms of pupil achievement.

- Few if any pupil benefits can be expected from reducing class size if teachers continue to use the same instructional methods and procedures in the smaller classes that they used in the larger classes.

- Some studies have found that even when teachers have small classes, many teachers do not take advantage of them to individualize instruction.

- Smaller classes appear to have a positive effect on pupil behavior in the elementary grades. At the secondary school level, some studies, but not others, have indicated that smaller classes influence student perceptions about their courses and their satisfaction with them.

- Researchers who have attempted to measure achievement gains in smaller classes over a relatively short period of time may not have allowed enough time for the desired changes to occur. Since some researchers have found that smaller classes must bring with them changes in instructional methods and teacher behavior before improved learning can take place, more longitudinal studies are needed to measure the effects of such possible changes over time.

- Opinion polls have consistently indicated that most teachers perceive large classes as a major factor negatively influencing teacher morale and job satisfaction plus the academic performance, personal development, and social development of pupils. Results of national opinion polls conducted among elementary school teachers indicate that about half of the teachers polled believed they could do their most effective teaching with a class containing 20 to 24 pupils and about one third believed they could do their best teaching with a class containing 25 to 29 pupils.

- Opinion polls show that the majority of the public perceives small classes as being of major importance to pupil achievement and progress.

- Class size is a major determinant of school system budgets. Even small system-wide changes of one or two pupils per class can have major impact on a school system's budget.

- In terms of pupil benefits, research findings fail to justify small overall reductions in class size or pupil-teacher ratio by a school board merely as a matter of general policy without definite pupil-benefit objectives for specific groups of pupils.

- Policy decisions pertaining to class size and pupil-teacher ratio involve factors that are complex, varied, and often emotionally charged. These require the weighing of the possible pupil benefits, the possible teacher benefits, the facilities utilized, the financial costs, and the possible political consequences.

APPENDIX C

Bibliography of Class Size Literature Reviews by Date of Publication

1979

Smith, Mary Lee and Gene V Glass. *Relationship of Class-Size to Classroom Processes, Teacher Satisfaction and Pupil Affect: A Meta-Analysis.* The Class Size and Instruction Project, Leonard S. Cahen, Principal Investigator. San Francisco, California: Far West Laboratory for Educational Research and Development (1855 Folsom Street, 94103), July 1979. 70 pp. ($5.50, fourth class mail; $7.00, first class mail).

1978

Class Size: A Summary of Research. Arlington, Virginia: Educational Research Service (1800 North Kent Street, 22209), 1978. 84 pp. ($12.50)

Glass, Gene V and Mary Lee Smith. *Meta-Analysis of Research on the Relationship of Class-Size and Achievement.* The Class Size and Instruction Project, Leonard S. Cahen, Principal Investigator. San Francisco, California: Far West Laboratory for Educational Research and Development (1855 Folsom Street, 94103), September 1978. 66 pp. ($5.50, fourth class mail; $7.00, first class mail).

Hess, Fritz. *Class Size and Student Achievement,* 1978. 23 pp. (ED 158 394)

Research Evidence Regarding the Impact of Class Size on Pupil Academic Achievement. Phoenix, Arizona: Maricopa County School Superintendent's Office (401 County Administration Building, 111 South Third Avenue, 85003), [1978]. 24 pp.

Thompson, Sydney. *Class Size.* School Management Digest. Series 1, No. 12. Burlingame, California: Association of California School Administrators (1575 Old Bayshore Highway, 94010), 1978. 38 pp. ($1.75 for ACSA members; $2.75 for nonmembers).

1977

Millard, Joseph E. *Small Classes? "What Research Says About Effective Small Classes and Possible Alternatives to Small Classes."* Ankeny, Iowa: Heartland Education Agency, January 1977. 14 pp. (ED 113 897)

1976

Ryan, Doris W. and T. Barr Greenfield. *Clarifying the Class Size Question: Evaluation and Synthesis of Studies Related to the Effects of Class Size, Pupil-Adult, and Pupil-Teacher Ratios.* Toronto, Ontario: Ontario Ministry of Education, 1976. 229 pp.

1975

Class Size and Teacher Load. Framingham, Massachusetts: New England School Development Council (85 Speen Street, 01701), April 1975. 63 pp. $3.00

Ryan, Doris W. and T. Barr Greenfield. *The Class Size Question: Development of Research Studies Related to the Effects of Class Size, Pupil/Adult, and Pupil/Teacher Ratios.* Toronto, Ontario: Ontario Ministry of Education, 1975. 342 pp.

1974

LaFleur, C. D., R. J. Sumner, and E. Witton. *Class Size Survey.* Canberra: Australian Government Publishing Service, 1974.

Schofield, Dee. *Class Size.* School Leadership Digest Series, Number Three. Arlington, Virginia: National Association of Elementary School Principals (1801 North Moore Street, 22209), 1974. 26 pp. (ED 093 031)

1973

Gajewsky, Stan. *Class Size: Review of the Literature and Selected Annotated Bibliography.* Reports in Education Number 2. A Monograph. Montreal, Quebec: McGill University, Faculty of Education, 1973. 58 pp. (ED 093 055)

1972

Shapson, Stanley M. *Optimum Class Size? A Review of the Literature.* Toronto, Ontario: Toronto Board of Education, Research Department, November 1972. 22 pp. (ED 070 757)

Templeton, Ian. *Class Size.* Educational Management Review Series Number 8. Eugene, Oregon: ERIC Clearinghouse on Educational Management, University of Oregon, August 1972. 8 pp. (ED 066 799)

1971

Reisert, John E. "Class Size," in *The Encyclopedia of Education,* vol. 2, pp. 157-160. (Edited by Lee C. Deighton.) New York, New York: Macmillan and The Free Press, 1971.

1970

Lindbloom, Dwight. *Class Size As It Affects Instructional Procedures and Educational Outcomes.* Minneapolis, Minnesota: Educational Research and Development Council of the Twin Cities Metropolitan Area, Inc., June 1970. 48 pp. (ED 059 532)

1969

Vincent, William S. "Class Size," in *Encyclopedia of Educational Research,* 4th ed., pp. 141-146. (Edited by Robert L. Ebel.) Toronto, Ontario: The Macmillan Company (Collier-Macmillan Ltd.), 1969.

1968

Class Size. Research Summary 1968-S1. Washington, D.C.: National Education Association, Research Division (1201 Sixteenth Street, N.W., 20036), 1968. 50 pp. (ED 032 614)

Sitkei, E. George. *The Effects of Class Size: A Review of the Research.* Research Studies Series, 1967-68. Los Angeles, California: Los Angeles County Superintendent of Schools, [1968]. 13 pp. (ED 043 124)

1964

Holland, Howard K. and Armand J. Galfo. *An Analysis of Research Concerning Class Size.* Research Contribution to Educational Planning, No. 11. Richmond, Virginia: State Department of Education, Division of Educational Research, November 1964.

1955

Ross, Donald H. and Bernard McKenna. *Class Size: The Multi-Million Dollar Question.* Study No. 11. New York, New York: Institute of Administrative Research, Teachers College, Columbia University, 1955. 24 pp.

1954

Blake, Howard E. "Class Size: A Summary of Selected Studies in Elementary and Secondary Public Schools." Ed.D. dissertation, Teachers College, Columbia University, 1954.

REFERENCES

The following is a numbered list of references cited in the text of this critique and related works consulted in its preparation. Where possible, addresses and prices are given to expedite the ordering of desired materials. Documents for which ERIC Document (ED) numbers are given can be ordered from the ERIC Document Reproduction Service, Computer Microfilm International Corporation, P.O. Box 190, Arlington, VA 22210. The price schedule for documents is as follows: *Hard Copy:* 1-25 pages, $1.82; 26-50 pages, $3.32; 51-75 pages, $4.82; 76-100 pages, $6.32. (Add $1.50 for each 25-page increment or fraction thereof.) *Microfiche:* 1-5 fiche, 83¢; 6 fiche, $1.00; 7 fiche, $1.17; 8 fiche, $1.34. (Add 17¢ for each additional fiche.)

1. "Argument for Smaller Class Sizes Raised Again by New Glass Study," *Education U.S.A.*, (September 10, 1979), pp. 9, 16.

2. "At Last, DATA on Class Size!" *PSEA Voice*, (February 26, 1979). (A publication of the Pennsylvania State Education Association.)

3. "Benefits of Smaller Classes Have Been Conclusively Documented," *Professional Developments*, (March 1979), pp. 5-6. (A publication of the Massachusetts Teachers Association.)

4. Burstein, Leigh. "Secondary Analysis: An Important Resource for Educational Research and Evaluation," *Educational Researcher*, 7 (May 1978), pp. 9-12.

5. Cahen, Leonard S. and Nikola N. Filby. "The Class Size/Achievement Issue: New Evidence and a Research Plan," *Phi Delta Kappan*, 60 (March 1979), pp. 492-495, 538.

6. *Class Size: A Summary of Research.* Arlington, Virginia: Educational Research Service (1800 North Kent Street, 22209), 1978. 84 pp. ($12.50)

7. *Class Size and Teacher Load.* Framingham, Massachusetts: New England School Development Council (85 Speen Street, 01701), April 1975. 63 pp. ($3.00)

8. "Class Size Is a 'Vital' Factor. . . ." *Education Summary*, (April 15, 1979), pp. 1-2.

9. Coleman, J. S. and Others. *Equality of Educational Opportunity.* Washington, D.C.: U.S. Government Printing Office, 1966.

10. Filby, Nikola, Leonard Cahen, Gail McCutcheon, and Diane Kyle. *What Happens in Smaller Classes? A Summary Report of a Field Study.* San Francisco, California: Far West Laboratory for Educational Research and Development (1855 Folsom Street, 94103), January 1980. 21 pp.

11. Glass, Gene V. "Integrating Findings: The Meta-Analysis of Research," pp. 351-379 in *Review of Research in Education 5.* (Edited by Lee S. Shulman.) Itasca, Illinois: F. E. Peacock Publishers, 1977.

12. _____. "Policy for the Unpredictable (Uncertainty Research and Policy)," *Educational Researcher*, 8 (October 1979), pp. 12-14.

13. _____. "Primary, Secondary, and Meta-Analysis of Research," *Educational Researcher*, 5 (November 1976), pp. 3-8.

14. _____ and Mary Lee Smith. "Meta-Analysis of Research on Class Size and Achievement," *Educational Evaluation and Policy Analysis*, 1 (January-February 1979), pp. 2-16.

15. _____. *Meta-Analysis of Research on the Relationship of Class-Size and Achievement*. San Francisco, California: Far West Laboratory for Educational Research and Development (1855 Folsom Street, 94103), May 1978. 93 pp.

16. _____. *Meta-Analysis of Research on the Relationship of Class-Size and Achievement*. The Class Size and Instruction Project, Leonard S. Cahen, Principal Investigator. San Francisco, California: Far West Laboratory for Educational Research and Development (1855 Folsom Street, 94103), September 1978. 66 pp. ($5.50, fourth class mail; $7.00, first class mail).

17. _____, Leonard S. Cahen, Mary Lee Smith, and Nikola N. Filby. "Class Size and Learning--New Interpretation of the Research Literature," *Today's Education*, 68 (April-May 1979), pp. 42-44.

18. "Glass Links Small Class Size to Increased Student Achievement," *Phi Delta Kappan*, 60 (February 1979), p. 411.

19. Hechinger, Fred M. "About Education: Smaller Classes Found to Produce Subtle Changes," *New York Times*, (April 10, 1979).

20. Henry, Neil. "Changes Voted in School Policy for Montgomery," *Washington Post*, (February 10, 1979), pp. B1-B2.

21. Hess, Fritz. *Class Size Revisited: Glass and Smith in Perspective*. East Syracuse, New York: East Syracuse-Minoa Central Schools, February 1979. 9 pp. (ED 172 402)

22. Light, Richard J. "Capitalizing on Variation: How Conflicting Research Findings Can Be Helpful for Policy," *Educational Researcher*, 8 (October 1979), pp. 7-11.

23. "Meta-Analysis of Research on Class Size," *ERS Bulletin*, 6 (February 1979), pp. 1-2.

24. Murphy, William F. *A Commentary on Class Size and the Glass-Smith Report*. Framingham, Massachusetts: New England School Development Council (85 Speen Street, 01701), June 1979. 5 pp.

25. Passarella, L. A. "The Relationship of Class Size to Teachers' Perceptions of Their Own Educational Effectiveness." Ed.D. dissertation, Rutgers University, 1977. (UMI Order No. 77-27954)

26. Shanker, Albert. "Where We Stand. Research Shows Smaller Is Better: Mayor Is Wrong--Class Size Matters," *New York Times*, (February 23, 1980), p. E9.

27. Shapson, S. M. and Others. "Results of an Experimental Study on the Effects of Class Size." Paper presented at the annual meeting of the American Educational Research Association, March 1978. (ED 151 985)

28. "Small Classes DO Make a Difference!" *PSEA Voice*, (February 26, 1979). (A publication of the Pennsylvania State Education Association.)

29. "Smaller Classes Better, Study Says," *Education Daily*, (October 29, 1979), p. 5.

30. Smith, Mary Lee and Gene V Glass. *Relationship of Class-Size to Classroom Processes, Teacher Satisfaction and Pupil Affect: A Meta-Analysis.* The Class Size and Instruction Project, Leonard S. Cahen, Principal Investigator. San Francisco, California: Far West Laboratory for Educational Research and Development (1855 Folsom Street, 94103), July 1979. 70 pp. ($5.50, fourth class mail; $7.00, first class mail.)

31. Walberg, Herbert J., Diane Schiller, and Geneva D. Haertel. "The Quiet Revolution in Educational Research," *Phi Delta Kappan,* 61 (November 1979), pp. 179-183.

32. Wright, Edgar N. and Others. *Effects of Class Size in the Junior Grades: A Study.* Toronto, Ontario: Ontario Ministry of Education, 1977. 130 pp. (Available from the Ontario Institute for Studies in Education, Publications Sales, 252 Bloor Street West, Toronto, Ontario M5S 1V6.)

Educational Research Service, Inc.
1800 North Kent Street
Arlington, Virginia 22209
Phone: (703) 243-2100