Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis
Veenman, Simon. **Review of Educational Research**. Washington: Winter 1995. Vol. 65, Iss. 4; pg. 319

Abstract (Summary)
Multigrade and multi-age classrooms, in which students from two or more grades are taught by one teacher at the same time, represent a significant phenomenon in schools. Veenman reviews the best evidence concerning the cognitive and noncognitive effects of the multigrade classroom and the multi-age classroom, in which children of different ageas are grouped together for educational and pedagogical benefits.

» Jump to indexing (document details)
Full Text (18893 words)

Multigrade classes are classes in which students from two or more grades are taught by one teacher in one room at the same time. Students in multigrade classes retain their respective grade-level assignments and their respective grade-specific curricula. Such classes are generally formed for administrative and economic reasons.

Some schools may deliberately mix both age and grade levels for educational reasons. The student is kept with the same teacher in the same class for a number of years, usually three. This practice is described as multi-age grouping. It is important to differentiate multigrade classes from multi-age classes. Multigrade classes are formed out of necessity; multi-age classes are formed deliberately for their perceived educational benefits.

A great deal of research has been devoted to the effects of multigrade and multi-age classes, but there are only a few comprehensive reviews on this topic. Based on a review of the studies of multigrade and multi-age grouping up to 1983, Pratt (1986) concluded that these types of grouping had no consistent effect on academic achievement and a generally benign effect on social and emotional development. Miller (1990) reviewed studies conducted up to 1986 and also concluded that multigrade and multi-age classrooms do not negatively affect students' academic performance or social relationships and attitudes. Both of these reviews were quite limited, however. Pratt did not provide a complete overview of the relevant studies, and in both reviews the number of statistically significant findings in favor of multigrade/multi-age classes or single-grade/single-age classes were simply counted. Little attention was paid to the particular form of multigrade/multi-age classes, the methodological quality of the studies, or the size of the effects. The criteria used for inclusion of studies were not indicated. Moreover, multigrade classes were not differentiated from multi-age classes. Both reviews included only studies conducted in English-speaking countries (primarily the United States).

The purpose of the present review is to highlight the reasons for the increasing number of multigrade classes, the concerns and perceptions about multigrade classes, and the claimed advantages of multi-age grouping. Moreover, the international research into the cognitive and noncognitive effects of multigrade and multi-age classes will be reviewed, and the implications for future research will be discussed.

The Prevalence of Multigrade Teaching

The multigrade classroom occupies a particularly significant position in schools in densely populated areas with shifting enrollment rates, in thinly populated areas, and in developing countries trying to improve the quality of education in rural communities. Since the mid 1970s, many schools in Western industrialized countries have been forced to discharge teachers because of financial cuts and declining student enrollments due to changing birth rates. As a consequence, these schools have had to set up multigrade classes (also called combination classes, mixed-age classes, split classes, double classes, and vertically-grouped classes). As a result of population shifts towards suburbs, schools in rapidly developing suburban areas may experience increases in enrollment while schools in older urban areas may experience declines. The administrative approach to dealing with these fluctuating enrollments has been to combine students from two or more consecutive grades to form a multigrade classroom with one teacher.

In the Netherlands, 53% of the elementary school teachers have multigrade classes (Commissie Evaluatie Basisonderwijs, 1994). In Switzerland, 23% of all classes are multigrade (Poglia & Strittmatter, 1983). In Germany, about 80,000 students attend schools with multigrade classes (Knoerzer, 1985). Of the 642 schools in England and Wales surveyed by Walsh, Dunne, Stoten, and Stewart (1984), 40% said that they had experienced an increase in multigrade group teaching as a result of declining enrollment. A further 15% said that declining enrollment might lead to an increase in the extent of multigrade teaching in the future. Almost half of the new teachers in England and Wales had their first posts in multigrade classes (Her Majesty's Inspectorate, 1982). In a sample from a suburban district in the United States (Arizona), approximately 17% of the students were in multigrade classes (Rule, 1983). Thus, demographic contraction and staff cuts appear to have caused increases in the number of multigrade classes in most Western countries over the past few decades.

As already noted, multigrading is a common phenomenon in rural areas where schools often consist of only one, two, or three teachers. One out of every seven classrooms in Canadian schools consists of two consecutive grades in one classroom. One out of every five students is enrolled in a multigrade classroom. It is interesting to note that in this large country more multigrade classrooms are found in urban than in rural districts (Gayfer, 1991). Over 85% of the elementary schools in Western Australia employ some form of multigrade grouping (Pratt & Treacy, 1986). One third of New Zealand's elementary schools are one- or two-teacher schools. In Scotland, a quarter of the elementary schools have fewer than 50 students. In Finland, 70% of the elementary school students are enrolled in schools with three or fewer teachers (Pietilae, 1978). In Portugal, 80% of the children attend schools with no more than two classrooms. In France, there are 11,000 one-teacher rural schools (Marshall, 198J). In Austria, about 25% of the elementary schools are one-, two-, or three-room schools (Noesterer, 1991). The United States currently has fewer than 1,000 one-room schools; as late as 1918, however, there were close to 200,000 one-room schools, which represented 71% of all public schools (Miller, 1990). This decline in the number of rural schools in the United States is due to population shifts, improved transportation, cost saving measures, and higher educational standards.

According to UNESCO (1989), the incidence of multigrade teaching in elementary schools is quite significant in Asian and Pacific countries. Countries such as the Republic of China and India have reported as many as 420,000 and 327,000 schools practicing multigrade teaching, respectively. Multigrade teaching occurs in about 20,000 schools in Indonesia and about 1,540 schools in Malaysia. A recent survey of elementary schools in Pakistan revealed that 58% of the elementary school teachers taught more than one grade at a time (Rowley, 1992). More multigrade schools are likely to be opened as the policy on equal educational opportunities is gradually implemented (UNESCO, 1989).

Advantages of Multi-Age Grouping

Multi-age student grouping is, of course, one characteristic of the nongraded school. However, it is also practiced by schools that do not label themselves as nongraded. As we will use the term, multi-age grouping occurs within the framework of a graded system, and students so grouped retain their grade labels. Schools that practice multi-age grouping claim that this classification is superior to single-age or horizontal grouping by stressing the advantages of the social and family-like structure of the group, in which students who are at least a year apart in age are placed in the same classroom for several years. Whereas multigrade classes are usually formed for administrative and financial reasons, multi-age classes are usually based on pedagogical and didactic motives. Many of the schools with multi-age classrooms have been influenced by the apparent advantages of "family groups" in nursery schools and by the principles underlying British infant schools, in which 5-, 6-, and 7-year-olds are taught in the same class (Mycock, 1966).

Advocates of multi-age grouping claim that it yields the following cognitive and noncognitive benefits.

(1) Students have a chance to form relationships with a wider variety of children than is possible in the traditional same-age classroom. This leads to a greater sense of belonging, support, security, and confidence.

(2) Teaching a diverse group of students demands individualized instruction.

(3) The development of a balanced personality is promoted by fostering the attitudes and qualities that enable students to live in a complex and changing social environment.

(4) The self-concepts of slower, older students are enhanced when they are asked to tutor younger students in their class.

(5) More secure teacher-student relationships may be established as the student remains with the same teacher for two or more years.

(6) Fewer anxieties may develop because the educational atmosphere is conducive not only to academic progress but also to social growth.

(7) Multi-age grouping provides younger students with the opportunity to observe, emulate, and imitate a wide range of behaviors; older students have the opportunity to assume responsibility for less mature and less knowledgeable students.

(8) Multi-age grouping invites cooperation and other forms of prosocial behavior and thus appears to minimize competitive pressures and the need for discipline.

(9) Students in the lower grade(s) can enrich their learning by attending to the material designed for the higher grade(s), while students in the higher grade(s) can profit from opportunities to review the material designed for the lower grade(s).

(10) Current concepts of cognitive development (e.g., the zone of proximal development and cognitive conflict) imply that children whose knowledge or abilities are similar but not identical can stimulate each other's thinking and cognitive growth.

(11) Finally, multi-age grouping relaxes the rigid curriculum with its age-graded expectations, which are inappropriate for a large number of students (Draisey, 1985; Ford, 1977; Franklin, 1967; Katz, Evangelou, & Hartman, 1993; Mycock, 1966).

Schools that are forced to create multigrade classes in order to meet enrollment needs often claim the same cognitive and noncognitive advantages as for multi-age grouping and thus make a virtue out of necessity.

Problems and Concerns Associated With Multigrade Grouping

The standardization of the single-grade format has created a biased attitude that single-grade classrooms are better than any other alternative. When a school switches from a single-grade organization to a multigrade organization out of necessity, both teachers and parents are usually not pleased. Responses from teachers in an early study of city schools in Connecticut by Knight (1938) showed that most teachers preferred a single-grade class over a multigrade class because the latter was felt to entail more planning, preparation, and work. Teachers in multigrade classes were generally found to teach the grades separately for mathematics and reading, and most of the teachers and principals felt, when considering students' educational progress, that multigrade classes were undesirable.

Canadian principals and teachers stated the following problems in connection with multigrade teaching: integration of the curricula, individualizing instruction, lack of time for adequate teaching of certain subjects, lack of time for the preparation of classroom materials, keeping up with marking, the need to attend to the one grade while instruction is being given to the other grade, the greater workload, lack of teacher time for

individual remediation, and lack of time to deal with parental concerns. When asked about the motives for organizing their schools using multigrade classes, most of the principals pointed to the unequal distribution of students. Almost 80% of the teachers responded that their multigrade classes had been assigned to them, and more than 80% of the respondents indicated that they had not had any training whatsoever in multigrade classroom teaching. Teachers named opportunities for students to learn from each other, for modeling, and for leadership as positive aspects of multigrade classes. About half of the principals felt that the optimal number of students in a multigrade class should be 20 or less, and that the enrollment should be lower in the lower grades because these students need greater individual attention. Half of the principals indicated that the parents were either negative or very negative about multigrade classes. These principals also indicated, however, that lack of understanding was the main reason for the parents' negative attitudes towards multigrade classes (Campbell, 1993; Daniel, 1988; Gayfer, 1991; Waraksa, 1989).

In a recent study of attitudes towards multigrade classes among students, parents, and teachers who had been actively involved in such classes in an urban public school district in Connecticut that generally adhered to the traditional single-grade approach, Walsh (1989) found the participants to agree on three points. First, they felt the greatest advantage of multigrade grouping was the opportunity for students at the lower level to advance academically due to exposure to the curriculum of students at the upper level. Second, they felt the major liability of multigrade grouping was lack of teacher time. And, third, no group of participants (students, teachers, or parents) would choose the multigrade structure again. The components of the "preferred model" for multigrade grouping that emerged from these suggestions included well-trained teachers, mature and independent students of average to high ability, and limited class size. The respondents also recommended only Grade 3/4 and Grade 4/5 combination classes for future multigrade grouping.

In a study by Pratt and Treacy (1986), teachers in Western Australia were actually surprised at being asked about the advantages of multigrade grouping. The following disadvantages were then pointed out by the teachers and principals: increased workload, more time required for the programming and preparation of materials, more time required for marking tests, not enough time for providing attention to individual students, and no opportunity to reflect on the teaching activities during the day. Rural communities were found to accept multigrade classes, whereas metropolitan communities were found to oppose multigrade classes or to accept them in only a very passive manner. The teachers were generally critical of teacher-training courses and claimed that the courses did not prepare them for teaching in multigrade classes. Pratt and Treacy's findings are largely confirmed by the results of Dutch and Swiss surveys (Casparis, Bernhard, & Heuberger, 1981; Inspectie Basisonderwijs, 1978; Mayer, 1981; Poglia & Strittmatter, 1983; Veenman, Lem, & Winkelmolen, 1985).

In sum, studies conducted at different times and in different countries have revealed some common problems and concerns associated with multigrade classes: lack of time for teaching the required content, a greater workload, lack of time for individual attention and remediation, lack of adequate classroom management skills, lack of adequate preparation during teacher training, inadequate materials, and parental concerns about the academic achievement of their children. These problems and concerns are stronger for teachers of multigrade classes in urban areas where the traditional single-grade approach is generally followed, and weaker for teachers of multigrade classes in sparsely populated areas where the small school is the norm.

In the next few years, financial cutbacks and concerns regarding efficiency will result in the closure of more than 1,000 small schools in the Netherlands. This closure is justified on the basis of cost savings and the assumption that bigger is better. It is argued that students in small schools with multigrade classes learn less than students in large schools with single-grade classes. In Britain, too, small schools are constantly being threatened and their viability has been increasingly challenged due to declining enrollments over the past few decades. One of the major issues in these debates is always whether education in a multigrade class can be effective or not (Hopkins & Ellis, 1991). In Germany and Austria, concerns about educational quality have already resulted in the closure of many small village schools with multigrade classrooms (Landschulen or Volksschulen) in favor of regional schools with single-grade classes (Mittelpunktschulen or Zentralschulen).

Given the advantages claimed by some to be associated with multigrade and multi-age grouping and the problems perceived by principals, teachers, and parents, the following question arises: What are the actual effects of multigrade or multi-age teaching on student learning? The purpose of the following review is to answer this question.

Review Methods

This review synthesizes the findings of research into the cognitive and noncognitive effects of multigrade/multi-age and single-grade/single-age classrooms in elementary school. The review method used is that of best-evidence synthesis as described by Slavin (1986), which combines elements of meta-analysis (Glass, McGaw, & Smith, 1981) with elements of narrative reviews. Briefly, a best-evidence synthesis requires locating all research on a given topic, establishing well-specified criteria of methodological adequacy and germaneness to the topic, and then reviewing this best evidence with attention to the substantive and methodological contributions of each study (Gutierrez & Slavin, 1992). Whenever possible, study outcomes are characterized in terms of effect size, the difference between the experimental and control means divided by the standard deviation for the control group. Details of the review procedure are described in the following sections.

Literature Search Procedures

An extensive search of the international literature was undertaken in order to identify potentially relevant studies. No restrictions were placed on the locations of the studies or the years of publication. The principal sources included ADION (Dutch Education Index), British Education Index (BEI), Dissertation Abstracts International, Education Resources Information Center (ERIC), European Documentation and Information System for Education (EUDISED), ForschungsInformationssystem Sozialwissenschaften (FORIS), FRANCIS (French-language literature), PsycLIT, Social Sciences Literature Information System (SOLIS; German-language literature). The following descriptors were used: multigrade classes, combination classes, multi-age classes, mixed-age classes, Mehrstufenklassen, Jahrgangskombinierte Klassen, classes a plusieurs degres, classes a cours multiples. In addition, the reference lists from earlier reviews and the primary studies themselves were used. Unpublished U.S. and Canadian doctoral dissertations were obtained from University Microfilms International (UMI) in Michigan. In the few cases where the unpublished documents were not available from ERIC or UMI, the institutions or authors were contacted directly. Every attempt was made to obtain a complete set of published and unpublished studies meeting the criteria outlined below. Of the final list of 56 studies included in this review, 33 were conducted in the United States, 4 in Canada, 3 in Germany, 3 in the United Kingdom, 3 in the Netherlands, 2 in Sweden, 2 in Finland, 2 in Australia, 1 in Colombia, 1 in Pakistan, 1 in Togo, and 1 in Burkina Faso.

Substantive Inclusion Criteria

Studies were included in the search if they involved the evaluation of the effects of multigrade or multi-age grouping. Multigrade and multi-age grouping were clearly distinguished in order to avoid an apples-and-oranges problem at the level of the independent variable. Multigrade grouping was taken to be an administrative measure for dealing with unequal class sizes or shifting student enrollments. Multi-age grouping was taken to be a pedagogical measure intended to enlarge the opportunities for interactions between younger and older students. In order to avoid an apples-and-oranges problem at the level of the dependent measures, the studies were also grouped as relevant to two major dependent variables: (a) academic or cognitive achievement and (b) noncognitive growth. The first area of relevance was further divided into the academic subjects addressed, for example, reading, language, mathematics, science, and social studies. The second area of relevance was further divided into (a) personal adjustment, (b) social adjustment, (c) self-concept, (d) attitudes towards school, and (e) motivation. Personal adjustment included variables such as feelings of belonging, absence of nervous symptoms, absence of antisocial tendencies, and absence of anxiety. An instrument frequently used in this domain is the California Test of Personality. Social adjustment included variables such as social adaptation to peers, school, class, and also to class and school transitions. Self-concept included variables such as self-

attitude, self-adequacy, and acceptance of self. Attitudes towards school referred to attitudes towards teachers, peers, school subjects, school social structure, and school climate. Motivation was related to achievement motivation or levels of aspiration. The dependent measures were analyzed separately in order to protect construct validity (Bangert-Drowns, 1986).

Studies of nongradedness were generally excluded from the present review. Although multi-age grouping is often the first step towards nongradedness, nongradedness today is also considered a philosophy of education that permeates the entire school organization and program (Anderson & Pavan, 1993). In short, the nongraded concept involves more than the vertical grouping of students for learning. Meta-analytic reviews of research into the effects of nongraded elementary schools on achievement have been conducted by Gutierrez and Slavin (1992) and by Anderson and Pavan (1993). The first of these studies was based on a comparison of effect sizes; the second used a vote-counting procedure. Of the studies reviewed by Gutierrez and Slavin, only those five that included descriptors such as multigrade, multi-age, combination class, or vertical grouping in their titles were selected for use in the present study. A study by Higgins (1980) was categorized as pertaining to multigrade grouping when close inspection revealed that the schools under study labeled themselves as "schools with combination classes." The remaining four studies (Burchyett, 1972; Givens, 1972; Mycock, 1966; Vogel & Bowers, 1972) were categorized as pertaining to multi-age grouping.

Methodological Inclusion Criteria

Only studies involving an explicit comparison of multigrade and multi-age classes with single-grade and single-age classes were selected for inclusion in the present review. In addition, inclusion required that a study meet a number of methodological criteria, which largely resembled those applied in earlier reviews on ability grouping and nongraded elementary schools by Slavin (1987, 1990) and Gutierrez and Slavin (1992). Following are the criteria used in the present review.

Experimental and control groups. All studies possessed both experimental (multigrade or multi-age) and control (single-grade or single-age) groups. Some studies did not meet this standard and were therefore excluded (Ford, 1980, 1981; Hoffman, 1981; Marturano, 1987; Mitchell & Zoffness, 1971; Way, 1958).

Standard measures. In all studies, standard measures of academic achievement or nonacademic achievement were used. Grades and report card scores were not included as achievement variables because of their subjective nature. Noncognitive variables were excluded if they were not based on some objective standard of measurement. Not all studies with noncognitive variables met this standard (e.g., Raberg, 1977, and a number of noncognitive measures in Mycock, 1966).

Comparability of samples. Ideally, initial comparability of the experimental and control samples was established by means of matching of schools or classes, or matching of individual students within classes or schools. Studies that used matching procedures and presented evidence that the groups were in fact initially equivalent are labeled in this study as matched studies with evidence of initial equality. Studies that used gain scores or analyses of covariance to control for initial differences between the experimental and control groups are labeled as matched studies lacking evidence of initial equality (with adjustment for pretest differences). Results of these studies should be interpreted cautiously because statistical adjustment for pretest differences cannot be assumed to control completely for their influences on posttests (see Reichardt, 1979). Studies that used matching of schools or classes with a posttest-only design (static-group comparison) are labeled in this study as matched studies lacking evidence of initial equality (without adjustment for pretest differences) and are listed at the bottom portions of the tables. The most obvious flaw of these studies is the absence of pretests, which raises the possibility that any posttest differences between the groups can be attributed either to a treatment effect or to selection differences between the different groups (Cook & Campbell, 1979). The results of such studies should be interpreted very cautiously. This category of studies was included for the sake of completeness.

Representativeness of the findings could also be established by means of random sampling procedures. Studies with random samples of schools or classes are listed in the tables in a separate category (studies using random samples). To ensure generalizability of findings and to control for threats of external validity, this category of studies used representative samples from the population of elementary schools with multigrade and single-grade classes. To control for internal validity in terms of selection factors (e.g., school size, locality, region), most studies used some form of stratified random sampling to add precision in ensuring that the samples contained the same proportional distribution of schools and classes on selected parameters as the population. Given the small number of elementary schools with multi-age classes, no studies with random samples were identified for multi-age grouping.

Only one part of one study was excluded on the basis of the comparability criterion. The reading achievement part of a study by Hoen (1972) was excluded because the results of the pretest and the IQ test showed that the experimental (n = 12) and control (n = 46) groups were not initially equivalent.

Duration of multigrade or multi-age grouping. In all of the included studies, the multigrade or multi-age grouping examined had existed for at least 1 year. All studies met this standard.

Normality of sample. All included studies involved samples of normal students in regular classes. This requirement excluded two studies involving gifted students (Runyon, 1962, 1963) and the mathematics achievement data from under-achievers in a study by Hoen (1972).

Teacher training. In all included studies, teachers in the experimental group had not been trained on the dependent measures. School improvement studies concerned with the training of teachers in multigrade or multi-age classes were excluded from the present review because the training effects might have obscure any grouping effects. This criterion excluded studies by Roelofs (1993) and Veenman and Raemaekers (1995).

Number of teachers. At least two experimental and two control teachers were involved in all of the studies included in this review. The purpose of this requirement was to minimize the degree to which teacher and class effects in small studies influenced study outcomes (cf. Slavin, 1987). This criterion resulted in the exclusion of two very small studies (Eames, 1989; MacDonald & Wurster, 1974).

Although the present review was predominantly concerned with elementary school students, two studies of the effects of multigrade or traditional single-grade elementary school classrooms on junior high school students were also included (Junell, 1971; Marsh, 1980).

No studies in which students were randomly assigned to experimental and control groups were found. A study conducted by Higgins (1980) had been categorized by Gutierrez and Slavin (1992) as a randomized study, but in the present study it was classified as a matched study lacking evidence of initial equality. Examination of the study showed that a principal had assigned students to combination and single-grade classes; other studies in this review suggest that principals often apply--either consciously or unconsciously--placement criteria such as independent work habits, cooperative attitudes, and ability in the assignment of students to different groups (see Adair, 1978; Brown & Martin, 1989; Spratt, 1986).

If a study met the above criteria but did not contain data that would allow for the computation of effect sizes, it was nevertheless considered in the present review. Such studies were discussed and included in the tables with an indication of the direction and the statistical significance of the observed differences.

Computation of Effect Sizes

In the literature on multigrade and multi-age grouping, single-grade or single-age grouping is almost always taken to be the control condition. This convention is followed in the present study, so that positive effect sizes are those favoring multigrade or multi-age classes and negative effect sizes are those favoring single-grade or

single-age classes. Whenever possible, effect sizes were computed in a manner similar to that employed by Slavin (1987, 1990) and Gutierrez and Slavin (1992). In general, effect sizes were computed as the difference between the experimental and control group means divided by the control group's standard deviation. The intention of this and other procedures was to put all effect sizes into the same metric, the unadjusted control group standard deviation. If pretest scores were available, posttest scores were adjusted using ANCOVA or raw gain scores. In studies that statistically adjusted posttest scores for one or more covariates, the difference in adjusted scores was divided by the unadjusted control group standard deviation. Similarly, in studies in which pretest data were provided, effect sizes were computed as the difference between experimental and control gain scores divided by the control group's posttest standard deviation. Whenever possible, denominators in the effect size computations were unadjusted individual-level posttest standard deviations, because use of aggregated data can spuriously inflate effect size estimates (cf. Slavin, 1987). For studies using class means (Adair, 1978; Brown & Martin, 1989; Carter, 1973; Her Majesty's Inspectorate, 1978), individual-level control group standard deviations were estimated by multiplying the class-level control group standard deviations by the square root of the average number of students per class. When the relevant means or standard deviations were missing, effect sizes were estimated from ts, Fs, ps, or other statistics, using procedures described by Glass et al. (1981) and Slavin (1987). For studies in which only the results of significance tests were reported and the p values were rounded to coarse approximations such as .05 or .01, the approximations were treated as accurate p values and thus provided conservative estimates of effect sizes.

For each study, effect sizes for the dependent measures were computed, along with an overall effect size estimate, which involved the average of the multiple effects. For certain purposes, effect sizes were pooled across studies. Whenever this was done, medians (and not means) were computed on all studies from which effect size estimates could be derived. According to Slavin (1987), pooling effect sizes within well-defined categories of studies can provide a useful summary of the size and direction of effects, but the pooled estimate should always be evaluated in light of the methodological quality and the consistency of the individual studies narratively described in the text.

An effect size of .20 is interpreted as a small or minimal effect, .50 is interpreted as a medium or moderate effect, and .80 is interpreted as a large or meaningful effect (Cohen, 1969). Asterisks by effect sizes indicate that the differences were statistically significant, according to the authors. When effect sizes could not be computed, the outcomes were classified as being in favor of multigrade/multi-age grouping (+), showing no difference (0), or being in favor of single-grade/single-age grouping (-), with asterisks if the differences were significant. This procedure is based on the work of Gutierrez and Slavin (1992).

Research on Multigrade and Multi-Age Grouping

In the following sections and tables, the results of research comparing the effects of multigrade/multi-age classes and single-grade/single-age classes on cognitive and/or noncognitive outcomes will be reviewed. The achievement or cognitive effects of multigrade versus single-grade classes will be considered first; the noncognitive effects of multigrade versus single-grade classes will be considered second; the cognitive effects of multi-age versus single-age classes will be considered third, and, finally, the noncognitive effects of multi-age versus single-age classes will be considered fourth. Within each section, the results of the matched studies with evidence of initial equality will be considered first, followed by the results of the studies using random sampling procedures and the matched studies lacking evidence of initial equality (with and without adjustments for pretest differences). Within each of these subsections, the larger studies are discussed first. In general, studies discussed earlier in each subsection of the text or listed earlier in a table can be considered higher in methodological quality than those discussed later in the subsection or listed later in the table (cf. Gutierrez & Slavin, 1992). Studies lacking evidence of initial equality will be summarized only briefly.

In discussions of the studies, the original terms used by the authors are used to denote the multigrade class. This means that equivalents such as combination class, double grade, split-grade class, mixed-age class, and vertically-grouped class may be used.

In most studies, information regarding the instructional practices in the multigrade or multi-age classrooms is not provided. In many of the studies, the number of years students spent in multigrade or multi-age classrooms is also not reported. When available, this information will be reported. When not available, the number of years spent in multigrade or multi-age classes is assumed to be one. For studies dealing with small (rural) schools, it is assumed that the students spent all of their school careers in multigrade classrooms.

Cognitive Effects of Multigrade Versus Single-Grade Classes

In this section, the results of studies dealing primarily with achievement effects of multigrade versus single-grade classes are considered. Studies dealing with both the cognitive and the noncognitive effects of multigrade classes are also considered in this section and will not be described in the section on the noncognitive effects of multigrade classes. Table 1 summarizes the research on the achievement effects of multigrade versus single-grade classes.(All Tables omitted)

Matched studies with evidence of initial equality. Nine studies compared multigrade classes to matched single-grade classes and presented evidence of initial comparability. In one of the largest matched-equivalent studies, Rule (1983) examined the effects of multigrade classes on student achievement in reading and mathematics in Grades 3 through 6 in Arizona. Each multigrade class was formed from students at two consecutive grade levels. Three grouping patterns were studied: multigrade classes, single-grade classes in multigrade schools, and single-grade classes not in multigrade schools. In addition, the achievement of the students in differing ability groups was analyzed. Three types of placement in multigrade classes were distinguished: high placement, average/high placement, and average placement. Multigrade classes with high-achieving students included students from the upper third in academic achievement, which was primarily a measure of reading achievement for both grades. For example, high-achieving second graders were placed with high-achieving third graders. A multigrade average/high class contained students from the middle and upper thirds in academic achievement in both grades. A multigrade average class combined average students from the lower grade with average students from the upper grade. The district under study was forced to use multigrade classrooms in order to economize and to equalize class loads. Overall, the multigrade classes did not appear to affect reading and mathematics achievement negatively (total ES = +.01). The average/high placement appeared to be best for all grades for reading and for Grades 4 to 6 for mathematics.

In a carefully matched study, Stone (1986) examined the possible effects of multigrade class placement on mathematics, reading, language, science, and social studies achievement in a large suburban school district in the United States. The multigrade classes were formed as a result of unequal enrollments and contained students from Grades 2 and 3. The results showed no significant differences between the multigrade students and the single-grade students in overall achievement (total ES = +.20).

One of the earliest studies of multigrade classes was conducted by Knight (1938) in Connecticut. In this study, multigrade classes were evaluated with respect to student achievement and any problems concerning the organization, administration, and teaching of double grades. The fourth-grade students in double third- and fourth-grade classes were matched with students in single fourth-grade classes; the fourth-grade students in double fourth- and fifth-grade classes were matched with students in single fourth-grade classes. The double-grade classes were generally organized as an economic means of handling the small number of students in the two consecutive grades. Slow students from the upper grade were generally combined with fast students from the lower grade. Results of this 1-year experiment indicated no differences between multigrade and single-grade classes (total ES = .00). The teachers in the double-grade classes generally taught mathematics and reading separately for each grade. The majority of teachers felt, when considering the student's educational progress, that double-grade classes were undesirable.

Veenman, Lem, Voeten, Winkelmolen, and Lassche (1986) collected both observational and achievement data for third- and fourth-grade students in combination and single-grade classes in the mid-eastern part of the Netherlands. The students and teachers in the combination classes were observed during 10 mathematics and 10

reading/language lessons. The students and teachers in the single-grade classes were observed during 5 mathematics and 5 reading/language lessons. The classroom observations occurred between pretest and posttest. No significant differences in test scores for the combination and single-grade classes were found (total ES = -.02). The observational data showed the time-on-task levels of the students in the combination classes to be about 6% lower than those of the students in the single-grade classes. Students in combination classes spent more time working individually than did pupils in single-grade classes. Most of the observed combination classes were organized so that the two grades (3 and 4) received instruction by turns. Little provision was made in the combination classes for individualized work or for similar-ability grouping; the students from the two grade levels were not further distinguished in terms of ability or attainment; seatwork involved students working alone on the same task as everyone else in the year group (Veenman, Voeten, & Lem, 1987).

Finley and Thompson (1963) examined the differences in achievement in basic academic subjects between students in multigrade classes and students in single-grade classes in rural California. Grades 3 and 5 were chosen for investigation, but the exact composition of the multigrade classes was not reported. The third graders had attended the school for the previous 2 years (as first and second graders); the fifth graders had attended the school for the previous 3 years (as second, third, and fourth graders). A multigrade school was defined as a school with four or fewer teachers. No significant differences in achievement scores for reading, language, or mathematics were found (total ES = -.06).

Lincoln (1981) assessed achievement differences among a sample of students who completed Grade 2 in Connecticut in the years 1977, 1978, and 1979. These students attended the same elementary school, where both the single-grade and multigrade patterns of organization had been in place for more than 7 years. The multigrade classrooms had been established to deal with fluctuating school enrollments. The school's classrooms had a differentiated instruction model whereby, for a large part of the school day, multiple instructional activities were going on at the same time. Students worked at instructional centers for language, writing, mathematics, science, arts and crafts. Much of their time was devoted to small-group activity. Both types of classrooms used the same basic reading program. Results of the study showed that reading achievement test scores did not vary significantly for children who spent Grades 1 and 2 in single-grade classrooms and those who spent Grades 1 and 2 in multigrade classrooms (ES = +.04). A small but significant difference favoring the multigrade classes was detected for the older students; older students in multigrade classrooms performed somewhat better than older students in single-grade classrooms (ES = +.19). Students with high intelligence scores in multigrade classrooms performed as well as students with high intelligence scores in single-grade classrooms (ES = +.08).

A study by Adams (1953) also showed no significant differences in the achievement of multigrade and single-grade students. These students were selected from combination and single-grade classes in California. Adams matched fifth-grade students who had been placed in combination fourth- and fifth-grade classes with students in regular fifth-grade classes. No student who had not attended the same school the preceding year was included. Students attended combination classes for at least 2 years. The differences in academic achievement between the combination and regular classes lacked statistical significance (total ES = +.13). A sociometric test was also administered in each class to examine the social structure of the class and the social adjustment of the students. The comparisons of different measures of integration or cohesiveness showed the percentage of friendship choices outside the class to be almost identical for the combination and regular classes; no consistent differences between the combination and regular classes with respect to the percentage of mutual friendship choices were found, either.

Furch-Krafft (1979) studied the influence of multigrade grouping and single-grade grouping on general scholastic achievement for fourth-grade students in rural areas outside of Freiburg, West Germany. No significant differences in overall achievement were found between the students in the multigrade classes and the students in the single-grade classes (ES = -.18).

In a study by Steinhofer (1980), children in combination and straight (traditional) classes in California were followed through kindergarten and first grade. A principal assigned the children to the two types of classes in the same manner as any principal assigns children to kindergarten. Instructional aides were employed in each of the combination kindergarten and first-grade classes, in each of the straight kindergarten classes, and in each of the straight first-grade classes. A scholastic achievement test was administered at three times: a pretest at the beginning of kindergarten, a posttest at the end of kindergarten, and a follow-up test at the end of first grade. In addition, a reading test was administered at the end of first grade. No significant differences between the straight and the combination classes were found (total ES = -.19). Differences between boys and girls were not found, either.

Studies using random samples. Achievement effects of multigrade and single-grade classes were compared in a total of 16 national and local studies using random-sampling techniques. In 5 studies, initial pretest differences between groups were dealt with statistically; 1 study corrected for background variables as proxies for the pretest. In 3 other studies, researchers compared the entire population of students from multigrade classes in a particular state or district with a random sample of students from single-grade classes. Characteristics of the samples are presented in Table 1.(All tables omitted)

Based on a representative sample of primary schools in England, Her Majesty's Inspectorate (1978) provided information on the organization of the particular schools, the extent to which the work of the students appeared to match their abilities, and the actual achievement of the students in reading and mathematics. Most of the vertically grouped or multigrade classes consisted of two or the age groups. The 11-year-old students in single-age classes received better reading and mathematics scores than the 11-year-olds in multigrade classes. The difference in reading scores for the 9-year-olds (who had no mathematics test) also favored single-age classes (total ES = -.26). The teachers were asked to identify groups of more capable, average, and less capable students within their classes. The relation between the work that the students could be expected to produce and the work that they actually produced was then assessed by the inspectors using classroom observations. When the degrees of match for the single-grade and multigrade classes we compared, definite superiority in relating the difficulty of the work to student capabilities at all levels was found for the 7- and 11-year-olds in single-age classes. No significant differences in the match assessments were found for the 9-year-olds.

For 3 consecutive years, Jokinen (1979) studied the achievement data from a national sample of Finnish elementary schools. The focus of the study was on centrally administered achievement tests in reading, language, Swedish and English as foreign languages, and mathematics (Grades 3 through 6). Students were divided into four groups: students from schools with 1-2 teachers, students from schools with 3 teachers, students from schools with 4-5 teachers, and students from schools with 6 or more teachers. For the purpose of this review, small schools with 1-2 teachers and schools with 3 teachers were taken to be schools with combined or multigrade classes; schools with 6 or more teachers were taken to be schools with only single-grade classes (1 teacher per class). In the small rural schools, the multigrade classes contained two or more grade levels. On the average, the achievement differences between the students in the multigrade classes and those in ordinary classes were near zero (total ES = +.03).

Jarousse and Mingat (1991a, 1991b, 1992) examined the language and mathematics achievement of second- and fifth-grade students in Togo and second-grade students in Burkina Faso. In these African countries, school instruction is in French. In both countries, the students in the multigrade classes were found to achieve significantly higher language and mathematics scores when compared to students at the same grade level in the single-grade classes. For Grade 2 in both Togo and Burkina Faso, an effect size of +.50 was found after controlling for student characteristics, teacher characteristics, class size, and school facilities. For the Grade 5 in Togo, an effect size of +.33 was found. The multigrade classes consisted of Grades 1 and 2 or Grades 5 and 6. The class sizes varied from 25 to 130 students in Too and from 23 to 150 students in Burkina Faso. For classes with more than 55 students, the negative effects of class size on student achievement were found to be more moderate for the schools with multigrade classes than for the schools with single-grade classes.

Gajadharsingh (1987) examined academic achievement in a large sample of students from multigrade and single-grade classrooms across the province of Saskatchewan, Canada. Only students enrolled in their respective grade levels (three, four, five, or six) and in one specific type of classroom (multigrade or single-grade) from the beginning of their school career up to the date of data collection were included in this study. The sample consisted of students from cities (urban areas), towns, villages, and hamlets (rural areas). No information was provided on the specific structure of the multigrade classes. The students in the multigrade classrooms were generally found to achieve significantly higher mean scores in language (vocabulary), reading, and mathematics when compared to students at the same grade level in single-grade classrooms (total ES = +.08). In hamlets, the achievement of multigrade students was significantly higher than that of single-grade students for language, reading, and mathematics. In cities, students in multigrade classrooms also performed significantly better than students in single-grade classes. However, students in single-grade classrooms located in towns achieved significantly higher mean scores in language than students in multigrade classrooms located in towns.

Brandsma (1993) studied the effectiveness of a random sample of Dutch primary schools and found an advantage of the single-grade classroom over the multigrade classroom. A multilevel approach with a pretest-posttest design was used to simultaneously analyze the data at the levels of school, class, and student. The achievement results for only the Grade 6 students were then studied. No information was provided on the exact structure of the combination classes. For the purpose of the present review, these data were reanalyzed using ANCOVA with preachievement, IQ, and socioeconomic status as covariates. Small but significant differences between the two groups were found for reading (ES = -.08) and language (ES = -.09), with the students in the single-grade classes performing better. No significant differences were found in mathematics performance (ES = -.04). The total effect size in the reanalyses of these data was -.07.

Rojas and Castillo (1988) studied the effects of multigrade versus single-grade instruction in rural Colombian schools and found an advantage for the multigrade classroom. Colombia, like many other developing countries, faces the dual challenge of improving the quality of education while increasing the availability of school education in rural areas. In the mid 1970s the Escuela Nueva, a rural school with one or two teachers providing 5 years of primary education, was introduced in a number of isolated rural areas with low population densities by the Ministry of Education to improve student learning. In the Escuela Nueva, students are promoted from one level to the next once they fulfill the minimum educational objectives of the first level. Special instructional materials such as student and teacher guides are provided. The curriculum is rural-oriented, and attempts are made to integrate the students and the schools into the communities by encouraging the parents to participate in school activities. In contrast, the traditional school, generally one classroom with one teacher, simply follows the national curriculum, does not provide special attention to slow learners, and does not stimulate the student with special materials. Along with Psacharopoulos, Rojas, and Velez (1993), Rojas and Castillo examined the cognitive achievement of third- and fifth-grade rural students in mathematics and Spanish, as well as their self-esteem. Schools with at least 5 years of Escuela Nueva experience were randomly selected for study. The mean scores on cognitive tests showed Escuela Nueva students to score higher than traditional students with the exception of mathematics among the fifth graders (total ES = +.22). The noncognitive test of self-esteem revealed no significant difference (ES = +.08). It should be noted that the Escuela Nueva is more than a school with a multigrade organizational pattern and that this form of teaching may explain only a small part of the findings in favor of the Escuela Nueva.

Spratt (1986) compared the reading scores of students enrolled in combination classes with those of students enrolled in regular classes in Virginia. The experimental group contained the following groupings: Grades 1 and 2, Grades 2 and 3, Grades 3 and 4, Grades 4 and 5, and Grades 5 and 6. Combination classes were formed as a result of unequal enrollments. Over one half of the principals of the schools with combination classes indicated they had chosen their better teachers to teach combination classes. The principals and teachers overwhelmingly (87%) said that they would not combine classes when given the choice. Overall, the levels of reading achievement in the combination classes proved to be no different from the level of reading achievement in the noncombination classes (total ES = +.28). Moreover, the achievement scores for the students in the upper grade

levels of the combination classes did not differ significantly from the scores for the students in the single-grade classes at the same grade levels (ES = +.27). Similarly, the scores for the students in the lower levels of the combination classes did not differ significantly from the scores for the students in the single-grade classes at the same grade levels (ES = +.36). The analysis of the data from surveys of the school principals showed that many of the better-than-average students were placed in the combination classrooms, which means that the weaker students may have been concentrated in the regular classrooms.

Marklund (1962, 1969) studied the achievement of students in sixth-grade classes in large and small schools in six Swedish counties. The students in the large schools attended single-grade classes while the students in the small schools attended double-grade classes (Grades 5 and 6) or four-grade classes (Grades 3 through 6). All of the schools were located in small towns or rural municipalities, and the distance between the student's home and school was kept constant in this study. The overall effect size for six academic subject areas was found to be -.18. The results showed that the students in the multigrade classes achieved equally to the students in the single-grade classes for class sizes up to 25. Beyond this figure, students in multigrade classes were found to perform at a significantly lower level than students in single-grade classes (ES = -.26).

Kral (1995) examined the effects of combination versus single-grade classes on the mathematics, language, and reading performance of second-, fourth-, and sixth-grade students in the Netherlands. The achievement gains of students in small schools (less than 110 students) versus large schools (more than 250 students) were of particular interest in this study. The small (urban and rural) elementary schools instructed their students in combination classes encompassing two or three grade levels whereas the large schools instructed their students in single-grade classes. As in the study by Brandsma (1993), a multilevel approach was used and, for the purposes of the present review, the data were reanalyzed using ANCOVA with preachievement, IQ, and socioeconomic status as covariates. No systematic differences were found between the combination and single-grade classes (total ES = -.06). Also, examination of teacher questionnaires and logs revealed no differences in the instructional time devoted to language, mathematics, and reading. The number of years spent in multigrade classes was not found to be associated with differences in achievement.

In a study of schools in rural Minnesota, Dreier (1949) held socioeconomic status, intellectual ability, and location of the school constant and found the ungraded or one-room school to be no better or worse than the graded school. The sixth graders from these two types of schools did not differ with respect to reading, language, mathematics, or spelling achievement.

In a representative sample of fourth-grade students in West Germany, Fippinger (1967) compared the achievement scores (a composite score based on tests of mathematics, reading, language, and social studies) for multigrade and single-grade schools. The students in the single-grade schools were found to score significantly higher than the students in the multigrade schools (ES = -.44). According to the author, this outcome can be explained by the fact that the single-grade schools had better students than did the multigrade schools. The single-grade schools we located in densely populated areas whereas the multigrade schools were situated in sparsely populated areas, and the socioeconomic status of the students in the (sub)urban areas was higher than that of the students in the rural areas. Unfortunately, neither socioeconomic status nor intelligence were controlled for in this study.

Nieminen (1979) reanalyzed Finnish achievement data collected in 1970 by the International Association for the Evaluation of Educational Achievement (IEA). The nationally representative Finnish sample contained third- and fourth-grade students (10-year-old children). The schools were divided into those with 1-2 teachers, those with 3 teachers, those with 4-5 teachers, and those with 6 or more teachers. For the purposes of the present review, the small schools--those with 1-2 or 3 teachers--were taken to be schools with combination classes; the schools with 6 or more teachers were taken to be schools with only single-grade classes. No significant differences in achievement could be discerned between the students from the schools with combination classes and the students from the schools with ordinary classes (total ES = +.02). There were also no differences in the students' attitudes towards school (ES = -.10) or motivation (ES = -.09).

Rowley (1992) performed a secondary analysis of the mathematics and science achievement data collected in Pakistan by Baloch (1990). A random sample of schools from four provinces in Pakistan was selected such that urban, rural, male, and female schools were proportionally included. The teachers were identified as single-grade teachers when they reported enrollments of only one class of students and as multigrade teachers when they reported enrollments of students in two or more grades. A significant difference in achievement was found in favor of the single-grade classes (total ES = -.36), and this difference did not appear to be influenced by the location of the school (urban versus rural) or the gender of the teacher. According to Rowley, there is considerable evidence showing the ineffectiveness of teacher training in Pakistan. A greater number of more highly trained teachers can be found in single-grade classes, which explains why students in these classes performed better. In addition, when teachers are assigned to teach in an area that is not their own ethnic or linguistic area, they may have to make use of student translators. When single-grade teachers used translators, they obtained significantly higher results than multigrade teachers. Teachers in multigrade classes made more use of student monitors to supervise the students who were not being taught and devoted significantly less time to direct instruction.

Harvey (1974) analyzed the effects of multigrade grouping on achievement, self-concept, and social-emotional development at the kindergarten level. All of the kindergartners in Grayson County (Virginia) were included in the study. The experimental group consisted of kindergartners in multigrade classes with first graders. The control group consisted of kindergartners in single-grade classes. All schools were located in rural areas, and the children in the multigrade classes had been grouped together to avoid the cost of a class with less than 20 children. No differences were found between the multigrade and single-grade groups (ES = -.07). The kindergartners in multigrade classes scored higher than the kindergartners in the single-grade classes on self-concept but not on social adjustment (total ES = +.44).

Zabolotney (1983) compared reading achievement and attitudes towards school among fourth graders in rural multigrade Seventh-Day Adventist schools and in rural single-grade public schools in Arkansas. The average number of students in the multigrade classrooms was 15; half of these classes consisted of Grades 1 through 4 and half of these classes had combinations ranging from Grades 3 and 4 to Grades 1 through 6. No significant differences were found between the two types of classrooms for reading achievement (ES = -.14) or attitudes towards school (ES = -.39).

Matched studies lacking evidence of initial equality. In 13 studies, multigrade classes were matched with single-grade classes without evidence of initial comparability. Six of these studies dealt statistically with initial differences among students. In seven studies, the researchers simply assumed that their samples were equivalent (posttest-only designs with nonequivalent groups). The characteristics of these 13 studies are described in Table 1.

Stimson (1991) studied the effects of multigrade classes on student achievement (Grades 3 through 6) in nine schools with year-round multitrack education in a large urban school district in California. Year-round education refers to the organization of the standard school days in such a way that instruction occurs throughout the year. Multitrack education allows the teachers to deal with over enrollment by having some of the students in attendance and others on vacation. The students in the single-grade classes were found to achieve at significantly higher levels than their counterparts in multigrade classes in reading, mathematics, and language (total ES = -.08). Also, the students in the upper level of a multigrade class were found to achieve at higher levels than the students in the lower level of a multigrade class (ES = +.11). According to the author, several things might explain the advantages of the single-grade classes in this study. First, in multitrack year-round schools, the parents select the track depending on the desired vacation time. Principals and teachers exert little influence on the composition of the multigrade class. Second, Quinlan, George, and Emmett (1987; cited in Stimson, 1991) found that multitrack year-round schools consistently performed below traditional-calendar schools in both reading and mathematics. From the results of this study, one may conclude that the multitrack year-round organization had a more profound (negative) effect on student achievement than did the multigrade classroom organization.

In their research program "Observational Research and Classroom Learning" (ORACLE), Galton and Simon (1980) studied the relative effectiveness of different teaching approaches across the main subject areas involved in primary school teaching in three local authority areas in Leicester, England. Most of the classrooms contained students from a single grade although some of the classes contained students from two age groups (7- and 8-year-olds, 8- and 9-year-olds, or 9- and 10-year-olds). No significant differences in the academic progress of the students were found. Observational data revealed slight differences in the teacher-student interactions and in the degree of student involvement.

A study of kindergarten and first-grade children in multigrade and single-grade classes was conducted by Adair (1978). This study was designed to compare the cognitive and noncognitive outcomes of single-grade and multigrade classrooms. It took place in the northeastern part of the United States. In some of the schools selected for the study, the principal was asked to select the "best" classroom(s) for inclusion. No significant differences were found for reading, language, or mathematical skills when the first graders in single-grade classes were compared to those in multigrade classes (total ES = -.06). With respect to the students' attitudes towards school, no differences were found between the kindergartners in the single-grade classrooms and those in the multigrade classrooms. The first graders in the single-grade classrooms, however, demonstrated a more positive attitude towards school than did the first graders in the multigrade classrooms (total ES = -.16)

Higgins (1980) compared the reading achievement of students in a traditional graded class with the reading achievement of students in a combination/ungraded class (Grades 3, 4, and 5) for three schools in Louisiana. The teachers in the combination/ungraded classes assigned the students to reading groups, but the instruction was highly individualized. No significant differences were found between the students in the graded classes and those in the combination classes (ES = +.05).

Rehwoldt and Hamilton (1957) reported the effects of newly created multigrade classes in a single California school. The multigrade classes were formed in units of three grades each (Grades 1 through 3, and Grades 4 through 6). The teachers were randomly assigned to multigrade or single-grade classes. The parents volunteered their children for the multigrade classes. Overall, no significant achievement differences were found for the multigrade versus regular-grade students (total ES = +.15). The findings with respect to personal and social adjustment also showed no differences. The attitudes towards school were found to be better for the multigrade students than for the single-grade students (total ES = +.32).

In a study conducted by Martens (1954) in an agricultural area of Iowa, students attended either one-room rural schools or graded town schools with one teacher per grade. Only those rural students who had spent all eight elementary grades in one-room schools and only those town students who had spent all eight elementary grades in a one-teacher-per-grade school were selected for study. The results showed the students educated in one-teacher-per-grade town schools to have significantly higher achievement scores than students educated in one-room rural schools (total ES = -.61). It must be noted that a highly significant difference in intellectual ability was found between the rural students and the town students; the latter demonstrated a higher level of ability.

Raberg (1976) examined the achievement of relatively isolated students who attended either multigrade or single-grade schools in the county of Vaesterbotten in Sweden. A child was considered relatively isolated if no other children of the same age lived within a kilometer of the child's home. Grades 3 and 6 were selected for examination. No significant differences were found between the relatively isolated students in multigrade classrooms and the relatively isolated students in single-grade classrooms (total ES = +.19). In Grade 6, however, the isolated students in multigrade classrooms were found to achieve at a much higher level in English as a foreign language than did those in single-grade classrooms (ES = +.42).

Brown and Martin (1989) examined the effects of multigrade classes in schools experiencing declines in enrollment in New Brunswick, Canada. The samples were selected from elementary schools that had both single-grade and multigrade classes. The students for this study were in fourth grade. Discussions with the principals and teachers showed that the children in the multigrade classes had been selected for maturity,

cooperativeness, and a willingness to work with minimal supervision. No significant differences were found for the effects of multigrade and single-grade grouping on overall achievement (ES = +.12).

Way (1981) explored the effects of multigrade grouping on the achievement and self-concept of students in the state of New York. The study included single-grade and multigrade classes of students in Grades 1 through 5. The multigrade classes included two or three ages together. No significant differences were found in the achievement performance of the students in the multigrade and single-grade classrooms (total ES = +.01). Similarly, no significant differences were found in the total self-concept scores (ES = +.17).

Knoerzer (1984, 1985) determined whether there were differences in cognitive and affective dimensions of schooling of third and fourth graders in Wuerttemberg, West Germany. Schools with multigrade classes were typically small schools in rural areas. No significant differences were found between the multigrade and single-grade classes in achievement. Similarly, no significant differences were found with regard to self-concept, attitudes towards school, personal and social adjustment, or motivation (total ES = +.28).

In California, Purl and Curtis (1970) explored the effects of student participation in combination classes on reading achievement, self-attitude, and personal and social adjustment. One combination class contained students from kindergarten through Grade 3, and the other contained students from Grades 1 through 3. The means for the combination and noncombination classes did not differ with respect to reading and self-concept. The students in the combination classes, however, were found to be less anxious about school and to perceive Black students more positively than the students in the noncombination classes.

Chace (1961) evaluated the academic, personality, and social development of students in Grades 3 through 6 in two schools in Tennessee. The school containing multigrade classes was a laboratory school at a teacher-training college located in a rural environment. One multigrade group included students from Grades 3 and 4; a second multigrade group included students from Grades 5 and 6; and a third multigrade group included students from Grades 3 to 6. The students in the third group were in this group primarily at the request of their parents. Multigrade grouping was found to offer no statistically significant academic advantage over single-grade grouping (total ES = +.20). Multigrade grouping offered a slight but consistently positive advantage with regard to personality and social development, which was indicated by statistically significant differences between the multigrade and single-grade groups on five of the eight subtests of the California Test of Personality (total ES = +.39).

Dodendorf (1983) compared students in a two-room rural schoolhouse in Nebraska with single-grade urban students. There were 19 students in the "lower room" (kindergarten through Grade 4) and 15 students in the "higher room" (Grades 5 through 8). Significant differences were found in favor of urban students in the area of social studies. No significant differences were found in the areas of language, mathematics, and science (total ES = -.24). Observation revealed that each grade in the two-room school met with the teacher, as a group, twice a day.

Summary of cognitive effects. Overall, the findings of research into the effects of multigrade and single-grade classes on achievement were found to be quite consistent. The median effect sizes for the three main categories of studies from which effect sizes could be estimated were: (a) ES = .00 for the 9 best-quality studies with evidence of initial equality, (b) ES = -.06 for the 15 best-quality studies using random samples, and (c) ES = +.03 for the 10 studies lacking evidence of initial equality. For the 34 studies for which effect sizes could be estimated, the median effect size was .00 (M = -.01).

Of the studies using random samples, four found significantly positive effects for multigrade classes (Gajadharsingh, 1987; Jarousse & Mingat, 1991b, 1992; Rojas & Castillo, 1988) and four found significantly negative effects (Brandsma, 1993; Fippinger, 1967; Her Majesty's Inspectorate, 1978; Rowley, 1992). Among the studies lacking evidence of initial equality, a negative effect was found by Stimson (1991) and Martens (1954). It should be emphasized that statistical significance is a product of both the treatment's effect and the

sample size. Conclusions from studies with larger samples ate more likely to be significant than conclusions from studies with smaller samples, even when the underlying treatment effect is small. A significant effect of about .08 in the studies by Gajadharsingh (1987), Brandsma (1993), and Stimson (1991) represents a difference between the multigrade and single-grade groups of about 8/100 of a standard deviation. Moreover, it should be noted that the educational setting in developing countries with multigrade classes differs largely from that in Western countries (Jarousse & Mingat, 1991a, 1991b, 1992; Rojas & Castillo, 1988; Rowley, 1992). Finally, in two of the earliest studies showing moderate negative effect sizes (Fippinger, 1967; Martens, 1954), the grouping effects were confounded with location effects (urban schools versus rural schools, and the urban schools had students of higher intelligence). Taking these considerations into account (together with a zero effect size from all 34 studies from which effect sizes could be estimated), one may conclude that students in multigrade classes learn as much as their counterparts in single-grade classes. Across a number of studies, the number of years spent in multigrade classes was also not found to be associated with differences in achievement.

Noncognitive Effects of Multigrade Versus Single-Grade Classes

In this section, the results of studies that focused exclusively on noncognitive effects of multigrade versus single-grade classes will be reviewed. As mentioned earlier, those studies that incorporated both cognitive and noncognitive effects are described in the section on cognitive effects. The research on the noncognitive effects of multigrade versus single-grade classes is summarized in Table 2.

Matched studies with evidence of initial equality. Three studies compared multigrade classes to matched control classes and presented evidence of initial comparability. Carter (1973) searched for differences in the attitudes towards self and the school among third and fifth graders in three school districts in Michigan. The multigrade classes consisted of students who had spent their entire school careers since kindergarten in multigrade classes. All of the designated schools had a composite achievement score that was above the national grade norm. The students in the multigrade classes obtained significantly higher scores than the students in the single-grade classes on measures of self-concept and attitudes towards schools (total ES = +.38).

Junell (1971) explored possible differences in the effects of a multigrade versus a traditionally graded elementary background on a variety of noncognitive variables among junior high school students. The junior high school, located in the state of Washington, received all of the students from a multigrade elementary school and all of the students from a regularly graded school. Only those students who had attended the elementary schools for a full 6 years were selected for the study. No significant differences were found, however, with regard to self-concept, attitudes towards school, or personality development (total ES = +.10).

Grove (1978) studied the effects of multigrade grouping and single-grade grouping on self-concepts and attitudes towards school among fourth, fifth, and sixth graders in parochial school systems in the states of Idaho, Washington, and Oregon. The schools in this parochial system were mainly located in sparsely populated areas. Only those teachers who had excellent principal and supervisor ratings were selected for the study. The schools under study were similar in terms of curriculum, teacher preparation, and community characteristics. No significant differences in the overall measures of self-concept and attitudes towards school were found (total ES = +.03).

Matched studies lacking evidence of initial equality. Three medium-size studies that lacked evidence of initial equality (without adjustment for pretest differences) reported no significant differences between multigrade and single-grade classes. In a study conducted in Western Australia, Pratt, Holub, and Ainsworth (1989) assessed whether students in multigrade classes (Grades 5 and 6) had different perceptions of their classroom environment than did students in single-grade classes. Students in multigrade classes were not found to possess a significantly more positive perception of their classroom environment than students in single-grade classes.

Pratt and Treacy (1986) compared the effects of multigrade versus single-grade classes on the attitudes of students in country and metropolitan schools in Western Australia. The multigrade students in this study were all in "year one-two" classes and did not differ significantly in their attitudes towards school from the students in the single-grade classes. Classroom observations revealed no major differences between the two class types. The general instructional formats in the multigrade classes tended to involve students working as an entire class or an entire grade level, with the teacher as the focus of attention. The students rarely interacted with each other in small groups or pairs. The time-on-task levels for the students in the multigrade classes in the country schools appeared to be about 6% higher than those for the students in the single-grade classes in the metropolitan schools--a difference that may be due to the smaller number of students in the country classes.

Hoen (1972) evaluated the influence of three multigrade classes on the reading skills, self-concept, motivation, and personal adjustment of fifth graders in a school in British Columbia, Canada. The three multigrade classes had been created 1 year prior to the study and contained an age range of 3 years. As mentioned before, the results of the cognitive part of this study are not considered in the present synthesis because pretests showed the reading achievement scores and IQs of the experimental and control groups to be unequal. No significant differences were found for the noncognitive variables (total ES = -.10).

Summary of noncognitive effects. In general, the research findings for the noncognitive effects of multigrade versus single-grade classes (see Table 2) resemble those for the cognitive effects. The students in the multigrade classes did not perform better or worse on noncognitive measures than did the students in the single-grade classes. The median effect sizes for the three main categories of studies for which effect sizes could be estimated had the following values: (a) ES = +.10 for the 3 best-quality studies with evidence of initial equality, (b) ES = -.01 for the 4 best-quality studies using random samples, and (c) ES = +.22 for the 6 studies lacking evidence of initial equality. For all 13 studies from which effect sizes could be estimated, the median effect size was +.10 (M = +.11).

Of the 17 studies reviewed, 5 reported significant noncognitive differences in favor of the multigrade classes (Chace, 1961; Harvey, 1974; Knoerzer, 1985; Purl & Curtis, 1970; Rehwoldt & Hamilton, 1957). However, the positive effects of multigrade classes on noncognitive measures in these studies were so small that they did not translate into higher academic achievement scores.

Cognitive Effects of Multi-Age Versus Single-Age Classes

In this section, the influence of multi-age versus single-age classes on cognitive and, in some cases, noncognitive variables will be considered. It should be recalled that multi-age classes are usually formed for deliberate educational and pedagogical reasons. Table 3 summarizes the research on the achievement effects of multi-age versus single-age classes.

Matched studies with evidence of initial equality. Of the 11 studies directed at the cognitive and noncognitive effects of multi- versus single-age grouping, only 2 studies presented evidence of initial comparability of the experimental and control groups. Mobley (1976) compared the effects of multi-age versus homogeneous-age grouping on instruction in an elementary school in Georgia. Half of the students in this school were already in multi-age classes, and the principal wanted to move to complete multi-age grouping in order to foster more individualized and better instruction in the future. The structure of the classroom and methods of instruction were left entirely up to the individual teachers in both types of classroom. The curriculum was the same for the two types of groups. Six homogeneous and seven multi-age classes had been established with pupils in the first, second, and third years of school. Students in the second and third years of school showed no significant differences in reading and mathematics achievement (total ES = .00). Students in the first year of school, however, showed significant differences in achievement; the multi-age group performed better. This finding has not been included in Table 3, however, because the pretest and the posttest were completely different. No significant differences were found between the groups on measures of self-concept (ES = +.18).

Givens (1972) evaluated the achievement and attitudes of fifth graders in two schools in Missouri. The demonstration school featured multi-age grouping of students, team-teaching, an open-space concept, and individualized instruction. Admission to this school was by application only, and the local universities contributed interns and teachers to the staff of the school. The results showed no significant differences in the overall achievement (reading, language, mathematics) of the students in the two schools (total ES = +.06). The results also showed the type of instructional program to produce no significant differences in the performance of the students at different levels of intelligence (high, average, and low). A significant difference was found, however, in general attitudes towards school (ES = +.58).

Marched studies lacking evidence of initial equality. Nine studies of the cognitive and noncognitive effects of multi-age classes lacked evidence of initial comparability. Five of these studies dealt statistically with initial differences among students, and four of the studies used posttest-only designs with nonequivalent groups without adjustment for pretest differences.

Schrankler (1976) examined the effect of multi-age grouping on student reading, mathematics, self-concepts, and attitudes towards school in a school in Minnesota. In one multi-age group, the age span was 7 years (ages 5 to 12 years); in the other multi-age group, the age spans were restricted to 2 or 3 years (ages 5 and 6; ages 7, 8, 9; and ages 9, 10, and 11). Traditional age groups were used as a control. On average, no differences in achievement were found. With respect to the students' self-concepts and attitudes towards school, only significant differences were reported.

Yerry and Henderson (1964) evaluated the effects of multi-age grouping on achievement and noncognitive variables such as anxiety and choice of friends in two schools in the state of New York. The students in Grades 1 to 6 in the multi-age classes (each combining two or three grades) received either small group or individualized instruction. No significant differences in the reading, mathematics, and language performance of the students in the multi-age and single-age classes were found (total ES = -.01). Also, no significant differences were found with regard to noncognitive variables (total ES = -.07).

Burchyett (1972) investigated the academic achievement, motivation, and self-concept of students in a multi-age, nongraded, team-teaching organization and students at similar grade levels in self-contained classrooms in Michigan. Each multi-age class contained students with an age range of 2 to 3 years. Instruction was provided on an individual basis by two or more teachers. The findings did not reveal significant differences with respect to either achievement (total ES = -.05) or self-concept and motivation (total ES = -.08).

Vogel and Bowers (1972) examined the effects of a multi-age nongraded school in Illinois on academic achievement, attitudes towards school, conceptual maturity, and performance during a standardized group problem-solving situation (kindergarten through Grade 6). Each class in the multi-age group encompassed an age range of 2 to 3 years. The students in traditional graded classes were found to have higher composite scores on the achievement test (ES = -.24). No systematic differences were found on the noncognitive measures (total ES = +.06).

The purpose of a study conducted by Marsh (1980) in Florida was to determine if students attending a middle school utilizing multi-age grouping achieved at academically higher levels than students attending a middle school where the grouping was according to chronological age. Students in the first school were assigned to a team(approximately 150 students and five teachers) and remained on that team for 3 years. Students in the second school were assigned to a different team of teachers each year. All of the students were in Grades 6, 7, and 8. Analysis of the fifth-grade data (pretest) revealed significant differences in reading, mathematics, science, and social studies. Eighth-grade scores (posttest) were subjected to analysis of covariance in an attempt to adjust for initial differences. Significant differences were found for most achievement tests in favor of the school with chronological age grouping (total ES = -.43). Given the many differences between the schools, however, these results should be interpreted with caution.

Milburn (1981) examined the effects of multi-age classes over a 5-year period in an experimental school in British Columbia, Canada. This school had five multi-age classes. In all of the classes, the students worked at different developmental levels. Each student remained in the same class with the same teacher for several years (usually 3). Milburn found little difference in the basic achievement of the students. The students in the experimental school, however, showed a more positive attitude towards school and a higher self-concept than their counterparts in the traditional grade groups.

Schroeder and Nott (1974) evaluated the effects of a multi-age grouping program one year after the program' s implementation in a school in Cincinnati, Ohio. The program was designed to allow each student to work at his or her individual skill level in each area of content. The multi-age group consisted of students in Grades 1 through 5, and the students in the fourth and fifth grades showed a greater amount of achievement than did their counterparts in the homogeneous groups. In addition, the students in the multi-age group were found to have a more positive attitude towards school.

Papay, Costello, Hedl, and Spielberger (1975) examined the relation between trait and state anxiety and performance on mathematical tests for disadvantaged first- and second-grade students in individualized multi-age and traditional classes in Dallas, Texas. No significant differences were found for mathematics achievement (ES = -.10). The students in the multi-age classes had lower twit anxiety scores than the students in the traditional classes. When matched for trait anxiety, the first graders showed no differences in state anxiety whereas the second graders in multi-age classes produced significantly lower state anxiety scores than did the second graders in traditional classes (total ES = +.29).

Mycock (1966) tested the assertion that at the infant stage vertical grouping is more beneficial than horizontal grouping. Four infant schools in Manchester, England were selected; two schools practiced multi-age grouping and two schools practiced single-age grouping. Children ranged between 5 and 7 years of age. Achievement tests were administered only to the 7-year-olds. No statistically significant differences were found between the multi-age group and the single-age group for reading or mathematics (total ES = +.13). Moreover, the results revealed no significant differences between the multi-age classes and the single-age classes on personal or social-adjustment variables (total ES = +.12).

Summary of cognitive effects. findings from this set of studies do not favor the multi-age classroom. In most of the studies, no significant differences were found. The median effect sizes for the two main categories of studies for which effect sizes could be estimated were: (a) ES = +.03 for the two best-quality studies with evidence of initial equality, and (b) ES = -.07 for the six studies lacking evidence of initial equality. For all eight studies from which effect sizes could be estimated, the median effect size was -.03 (M = -.08). In the study by Marsh (1980), the largest significant achievement differences were found in favor of the single-grade classes, although significant pretest differences bring these findings into question. In terms of academic achievement, multi-age classes appear o be generally equivalent to single-age classes.

Noncognitive Effects of Multi-Age Versus Single-Age Classes

The research on the noncognitive effects of multi-age versus single-age classes is summarized in Table 4. As mentioned before, those studies that incorporated both cognitive and noncognitive effects are described in the previous section on the cognitive effects of multi-age versus single-age classes.

Matched studies lacking evidence of initial equality. Only one study was found to be explicitly directed at the noncognitive effects of multi- versus single-age classes. Hammack (1974) studied the relation between the single-age and the multi-age classroom and the self-concepts of 3-, 4-, and 5-year-old children in Texas. In both the comparison and the experimental groups, the classrooms were arranged in interest centers incorporating an informal approach to teaching. The classrooms were comparable in physical appearance, curriculum, and daily routine. The self-concept scores in the multi-age groups were found to be significantly higher than the self-concept scores in the single-age groups (ES = +.33).

Summary of noncognitive effects. The median effect sizes for the two main categories of studies for which effect sizes could be estimated were: (a) ES = +.38 for the two best-quality studies with evidence of initial equality, and (b) ES = +.09 for the six studies lacking evidence of initial equality. The finding in the first category of studies can be largely attributed to the results of the study by Givens (1972), who found a medium-size effect in favor of the multi-age classes for student attitudes towards school (ES = +.58). For all eight studies from which effect sizes could be estimated, the median effect size was +.15 (M = +.18). The findings on self-concept and attitudes towards school suggest a small positive effect for students in multi-age classes.

Results and Discussion

Schools with multigrade/multi-age classes occupy a unique place in the history of education. The multigrade/multi-age school was the dominant model of education until the arrival of the industrial revolution and urbanization. Single-grade schools gradually emerged during the nineteenth century as a means of managing students by organizing them into age divisions or grades. Schools with multigrade classes are still an important organizational form in many suburban and rural areas throughout the world, however. Such schools are often the only source of education in rural areas. Multigrade schools also offer small towns and villages an alternative to uneconomical single-grade schools and help preserve the identity of the local communities. In developing countries, multigrade schools are seen as an efficient means of providing access to primary education under severe budgetary and work force constraints (Thomas & Shaw, 1992). In urban areas, multigrade classes have been established in response to declining student enrollments and/or staff reductions.

Parents and teachers often oppose multigrade classes, particularly in urban areas where schools have been forced to establish multigrade grouping in response to declining enrollments. The parents generally object because they believe their children will learn less in multigrade classes. For parents living in rural areas, the only viable option is often a multigrade setting. These parents seek assurance that multigrade classes will not hinder their children's learning. The teachers generally object to the additional workload and difficulty of catering to the wide range of ages and abilities found in a multigrade class. Researchers sometimes support this opposition by interpreting their results rather carelessly. For example, Stimson (1992) recommended that schools "avoid use of combination classes" (p. 17) and "make the avoidance of combination classes a number one priority in your class size reduction plan" (p. 18). The results of Stimson's research, however, do not justify such strong claims.

A review of the research into the effects of the multigrade/multi-age classroom is particularly necessary today in light of the growing prevalence of such classrooms and the lack of consensus with respect to this educational form. In this article, the best evidence with regard to the cognitive and noncognitive effects of multigrade/multi-age versus single-grade/single-age classes in elementary schools was reviewed. Two principal grouping plans for combining grades into classes were examined: multigrade grouping and multi-age grouping. Multigrade grouping was defined as an administrative device used to cope with declining student enrollments or uneven class sizes. It was considered an administrative necessity. Multi-age grouping was defined as the deliberate grouping of children with different ages into the same classroom for educational and pedagogical reasons.

Table 5 summarizes the cognitive and noncognitive outcomes of the 56 studies that met the inclusion standards for this review. The upper part of Table 5 summarizes the outcomes of the 45 studies concerning multigrade grouping. As can be seen, the students in the multigrade classes do not appear to learn more or less than their counterparts in the single-grade classes. No consistent differences were found with respect to reading, mathematics, language, or composite scores. In 28 of the 38 studies of the effects of multigrade versus single-grade classes on total achievement, significant differences were also not found. In 4 of these studies, significant positive effects were found; in 6 of the studies, significant negative effects were found. The median effect size across the 34 studies for which effect sizes could be computed was essentially zero (ES = .00). The same pattern generally emerges for the studies with noncognitive measures. The median effect size for all 13 studies from which effect sizes could be estimated approached zero (ES = +.10). If differences were found, they were very small. In 12 of the 17 studies of the effects of multigrade versus single-grade classes on noncognitive outcomes,

no significant differences were found; in 5 of these studies significant differences were found in favor of the multigrade classes. This suggests that in affective areas such as attitudes towards school, self-concept, and personal and social adjustment, students are sometimes better off in multigrade classes than in single-grade classes.

The lower part of Table 5 summarizes the outcomes of the 11 studies concerning multi-age grouping. As was the case for multigrade grouping, the students in the multi-age classes did not learn more or less than the students in the single-age classes. The median effect size for the 8 studies for which effect sizes could be computed was again essentially zero (ES = -.03). In 9 of the 11 studies of the effects of multi- versus single-age grouping, no significant achievement differences were found. With regard to the noncognitive outcomes, the number of significant positive findings exceeded the number of studies in which no significant differences were found. Students in the multi-age classes tended to score higher on attitudes towards school, personal adjustment, and self-concept. Once again, however, the affective and psychosocial differences between the students in the multi-age classes and those in the single-age classes proved to be very small. The median effect size across the 8 studies for which effect sizes could be computed approached zero (ES = +.15). In conclusion, parents, teachers, and administrators need not worry about the academic progress or social-emotional adjustment of students in multigrade or multi-age classes. These classes are simply no worse, and simply no better, than single-grade or single-age classes.

Table 6 summarizes the outcomes of the studies reviewed here in terms of the locality and socioeconomic background of the schools under study. No differences in achievement were found between students in multigrade classes and students in single-grade classes, in either (sub)urban areas (ES = +.01) or rural areas (ES = -.10). With regard to noncognitive effects, a small positive effect size was found in favor of the multigrade classroom in rural areas (ES = +.28).

Multi-age classes were found only in (sub)urban areas. (In rural areas, schools often have no choice other than to form multigrade classes.) No significant cognitive or noncognitive differences were found between students in multi-age classes and those in single-age classes in (sub)urban areas (ES = -.03 and ES = +.15 for cognitive and noncognitive measures, respectively). In 5 of 10 studies, students in multi-age classes in (sub)urban areas tended to score higher on noncognitive measures.

In the lower part of Table 6, the socioeconomic backgrounds of the schools are outlined and not found to affect the cognitive or noncognitive variables. A small positive noncognitive effect was found for the multigrade classes with students from the upper/middle socioeconomic backgrounds (ES = +.25). In only one study were the effects of multi-age classes with students from lower socioeconomic backgrounds examined. The lack of studies of multi-age classes with students from lower socioeconomic backgrounds is probably due to the fact that multi-age grouping has been conceived predominantly for students with upper- or middle-class backgrounds.

Finally, the differences in achievement for the multigrade versus single-grade settings were not found to vary significantly across the different grade levels. The effect sizes were close to zero for each level: for the lower grades (kindergarten through Grade 2), ES = +.04; for the intermediate grades (Grades 3-4), ES = -.0l; and for the higher grades (Grades 5-6), ES = +.05. Overall, the same pattern was found for the noncognitive variables and for the comparison of multi-age versus single-age settings.

Several factors may help explain why student learning in multigrade or multi-age classes does not differ from student learning in single-grade or single-age classes. First, it is unlikely that the grouping alone will affect student learning. Successful learning is less dependent on organizational strategies than on the quality of the instructional practices. At present, the instructional practices found in multigrade or multi-age classes are poorly understood. In fact, most studies provide no information whatsoever on the instructional practices employed in the classroom. Future research should examine not only the effects of different forms of organizational grouping but also the processes by which these effects are brought about. Those studies that do provide information on

the instructional processes in multigrade classes suggest that the most popular method is to teach a lesson to one grade while the other grade works on follow-up activities to previous instruction (individual seatwork). The teachers' time is usually divided between two or more groups, instructed separately. Single-grade-level instruction is emphasized. The students in multigrade classes are not necessarily engaged in learning from each other. Each student essentially works on his or her own in a group setting. Learning is not a collaborative accomplishment. The possibilities for the effective grouping of students in order to reduce student heterogeneity and foster the appropriate pace and level of instruction for each individual are simply not exploited. Students are rarely regrouped for instruction across grade lines. The studies by Slavin (1987) and Gutierrez and Slavin (1992), however, have shown that cross-grade grouping can result in consistent positive achievement effects. Grouping students across grade or age lines may allow teachers to reduce the number of reading and mathematics groups within a class, thus reducing the need for independent seatwork and follow-up. Students generally spend more time on independent seatwork in multigrade classes than in single-grade classes. Moreover, students' time-on-task levels have been shown to be lower during independent seatwork than during teacher-directed instruction (Evertson, 1989). Grouping students across grades provides teachers more time for direct instruction. Another factor in favor of cross-grade grouping is the likelihood that flexible cross-grade grouping allows teachers to adapt instruction to the individual needs of each student while still delivering instruction to the entire group. Using ineffective grouping arrangements and relying on single-grade instruction make the multigrade class no different from the single-grade class and thus minimize the potential positive effects of multigrade grouping. Incorporating a great deal of individualization might also reduce the effectiveness of the multi-age classroom. According to Gutierrez and Slavin (1992), individualized instruction, learning stations, learning activity packets, and other individualized or small group activities reduce direct instruction time with little corresponding increase in appropriateness of instruction to individual needs.

Research on effective instruction has consistently shown that student learning is enhanced by direct instruction from teachers, as opposed to extensive reliance on individualization, seatwork, and written materials (Brophy & Good, 1986; Rosenshine & Stevens, 1986). The move to greater individualization in multi-age classes at the expense of direct instruction might explain why student learning in the multi-age classes reviewed in this article does not differ from that in single-age classes. The findings in this review also correspond to those of Gutierrez and Slavin (1992), who evaluated nongraded programs incorporating individualized instruction.

A second factor that may help explain why student learning in multigrade classes does not differ from student learning in single-grade classes is bias in the composition of multigrade classes. Several studies suggest that when principals and teachers select students for placement in multigrade classes, they employ conscious criteria (e.g.; Adair, 1978; Bennett, O'Hare, & Lee, 1983; Brown & Martin, 1989; Carleton Board of Education, 1991; Gayfer, 1991; Spratt, 1986). Among the frequently mentioned selection criteria are independent work habits, cooperation, and no behavioral or emotional difficulties. These selection criteria are applied mostly in schools located in urban areas. They of course lead to nonequivalent samples such that the multigrade classes have more independent or autonomous students. To what extent the schools with multigrade classes in the present review selected students on these criteria is unknown. The effect of student selection criteria is speculative but may account for the absence of differences in student outcomes in some of the reviewed studies. On the other hand, most of the studies in rural areas, where student selection is simply not an option, also found no differences in student learning between the two grouping arrangements.

A third factor that may help explain the absence of differences is that teachers in multigrade classes are ill-prepared to teach two or more grades at the same time and may no have teaching materials that ate adequately suited to multigrade teaching at their disposal. Teachers often reported no special training for multigrade teaching, and the materials teachers used rarely addressed the issue of multiple grades. Inadequate training and teaching materials may strengthen the negative attitudes of teachers towards multigrade teaching and help maintain the single-grade class as the norm for educational practice.

A fourth factor that may help explain the absence of differences between multigrade and single-grade classes is the fact that most of the teachers indicated that multigrade classes impose a greater workload, require more

preparation time, and demand better classroom management skills. Somehow teachers in multigrade classes succeed in allowing their students to learn as much as students in single-grade classes. Given the demanding nature of multigrade classes, however, the teachers have little opportunity or energy for the use of potentially more effective grouping arrangements and therefore stick to the same practices as in single-grade classes.

At present, we do not know which combinations of grades are most effective. In their research on the use of mixed-age classes in England, Bennett et al. (1983) found more than 200 variations in types of combination classes. Some studies suggest that the lower grades are best suited for combinations (e.g., Gayfer, 1991), whereas others recommend only the Grade 3/4 and Grade 4/5 combinations (e.g., Walsh, 1989) or the higher grades (e.g., Daniel, 1988). Research on this topic is greatly needed.

We also know little about the optimal class size for multigrade teaching. When asked, teachers and principals state that between 20 and 25 students should be the optimum (Gayfer, 1991). Marklund (1962) found that sixth-grade students in multigrade classes and their counterparts in single-grade classes achieved equally well for class sizes up to 26 students. Beyond this figure, students attending single-grade classes produced higher achievement scores. In their study of classes in two African countries, Jarousse and Mingat (1991a, 1991b, 1992) found the negative effects of a large class size to be more moderate in the multigrade setting than in the single-grade setting. In fact, the benefits of multigrade teaching outweighed the negative effects of class size for classes of more than 55 students. In other words, the optimal multigrade class size has yet to be determined and should be examined separately for developed and developing countries.

In developing countries, multigrade teaching has proved to be an efficient means of providing education for difficult-to-reach populations (Black, 1982; Thomas & Shaw, 1992). Nevertheless, multigrade teaching is often seen as a second-class solution elicited by growing resource constraints and adverse demographic conditions. In many developing countries it is assumed that a teacher familiar with the instruction of a single-grade classroom can function successfully in a multigrade classroom without additional training in instructional techniques that are essential for effective multigrade teaching--self-directed learning, peer tutoring, careful lesson planning, variation in methods of instructional delivery--and without appropriate teaching materials (Thomas & Shaw, 1992), Studies from developing countries show mixed results for multigrade programs. Jarousse and Mingat (1991a, 1991b, 1992) report that multigrade students in Togo and Burkina Faso performed better than single-grade students and attributed this to the effects of peer tutoring, independent work, and a variety of presentation methods. The study by Rojas and Castillo (1988) shows that multigrade schools can be effective with the commitment of teachers, the use of special instructional materials, the provision of in-service activities, and the involvement of parents in the educational process. The multigrade structure with dedicated people may be effective. In contrast to the studies conducted in Africa and Colombia, for example, multigrade students in Pakistan were found to perform worse on achievement tests than their single-grade counterparts (Baloch, 1990; Rowley, 1992). In Pakistan, the teachers were not supported by a training program and were not provided with appropriate texts and teaching materials, which suggests that much more research is needed on the effects of multigrade teaching in developing countries and on the conditions under which multigrade teaching can be successfully implemented and adapted to local conditions.

As noted earlier, one characteristic of the nongraded school is multi-age grouping. In the present study, the different forms of multi-age grouping produced no consistent positive achievement effects. In their review of the effects of nongraded organization on achievement, however, Gutierrez and Slavin (1992) reported positive achievement effects for programs involving just one subject area, known as Joplin-like programs, and programs involving multiple subjects areas, called comprehensive programs. In the Joplin-like programs, students are grouped across age lines for a single subject (usually reading). In comprehensive programs, students are grouped across age lines for many subjects. The features that appear to be important to almost all of the successful nongraded programs were found to be flexibility of student grouping, frequent assessment of mastery, increased amounts of teaching time for homogeneous instructional groups, subject areas organized by levels, and use of texts written in accordance with those levels. Gutierrez and Slavin also examined nongraded programs that incorporated a great deal of individualized instruction (and correspondingly less teacher-directed

instruction) and found them to be consistently less associated with achievement gain than other nongraded programs. Extensive reliance on individualization, seatwork, and written materials does not appear to enhance student learning. A final group of studies for which Gutierrez and Slavin could not determine the nature of the nongraded program was also found to be less associated with achievement gain than the Joplin-like and comprehensive programs.

The 11 studies of multi-age grouping examined in the present study had several characteristics in common with the nongraded programs examined by Gutierrez and Slavin (1992) that incorporated a great deal of individualized instruction. Almost all of the studies of nongraded programs incorporating individualized instruction in the Gutierrez and Slavin review and 9 of the 11 studies of multi-age grouping in the present review were published in the 1960s and early 1970s, when individualization was gaining popularity. As shown by Gutierrez and Slavin, incorporating a great deal of individualization might have reduced the effectiveness of a nongraded school organization. In 2 of the multi-age studies in the present review, the instructional practices were poorly described, which means that the exact differences between the experimental and control groups were unclear. The remaining 9 multi-age studies may have produced no consistent positive achievement precisely because of their emphasis on individualized instruction. Finally, it should be noted that only 2 of the 11 multi-age studies provided evidence of initial equality.

Teachers working with small groups of students in multigrade or multi-age classrooms might improve the productivity of these groups by using forms of cooperative learning. Cooperative learning is defined by Cohen (1994) as students working together in a group small enough for each student to participate in a clearly assigned collective task. Students are expected to carry out the task without direct and immediate supervision by the teacher. Cooperative learning has gained increasing acceptance as a strategy for fostering learning gains, developing higher-order thinking, and encouraging prosocial behavior. Cooperative learning has also gained acceptance as a way to manage academic heterogeneity in classrooms with a wide range of basic achievement skills. Most models of cooperative learning in fact advocate the use of heterogeneous groups, and there is considerable research showing that cooperative learning in heterogeneous groups is particularly beneficial for low-achieving students (Cohen, 1994). Cooperative learning was not used, however, in the studies of multigrade end multi-age classes reviewed here. Future research should therefore explore the potential benefits of cooperative learning in multigrade and multi-age school settings.

Yet another instructional approach that deserves attention in the context of multigrade or multi-age teaching is reciprocal teaching, which also encourages students to provide instructional support for each other. Designed to improve reading comprehension (Palincsar & Brown, 1984; Rosenshine & Meister, 1994), reciprocal teaching is a collaborative learning procedure used to support the discussion of a text' s meaning and thus the development of comprehension skills. The instruction occurs primarily in the context of a dialogue between the teacher and the students and between the students themselves. All members of the group take turns leading the discussion. When the members are not leading the discussion, they support it by providing additional explanation of the content, requesting clarification, and solving any misunderstandings. Though the teacher initially assumes major responsibility for this dialogue and stimulates the participation of each student, the goal is to gradually transfer control of the discussion to the students (Palincsar, Stevens, Gavelek, 1989).

On still another front, Miller (1989, 1991) identified six key variables for the success of multigrade/multi-age teaching: (a) classroom organization, (b) classroom management and discipline,(c) instructional organization and curriculum, (d) instructional delivery and grouping, (e) self-directed learning, and (f) peer tutoring. In the multigrade/multi-age classroom, more time must and should be spent on the organization and planning of the instructional process. The teacher cannot be with every student at the same time and should therefore share the instructional responsibilities with the students within a context of clear rules and routines. The students should know what to expect and they should learn to help one another and themselves. Interdependence and independence should both be promoted. The teacher should emphasize the similarities among the different grades and thus conserve valuable instructional time. In multigrade/multi-age classes the students should serve as teachers to other students within and across differing grade levels. These key variables should be addressed

in preservice and in-service training programs in order to make the multigrade/multi-age classroom as effective as it can be. Teacher training institutions should acknowledge that the multigrade/multi-age class is a present and future reality.

References

Adair, J. H. (1978). An attitude and achievement comparison between kindergarten and first grade children in multi and single grade classes. Unpublished doctoral dissertation, Boston College.

Adams, J. J. (1953). Achievement and social adjustment of pupils in combination classes enrolling pupils of more than one grade level. Journal of Educational Research, 47, 151-155.

Anderson, R. H., & Pavan, B. N. (1993). Nongradedness: Helping it to happen. Lancaster, PA: Technomic.

Baloch, I. S. (1990). Impact of multi grade and single grade teaching on students achievement in Pakistan. Paper presented at the BRIDGES Training Workshop on Analysis of Survey Data, Academy of Educational Planning and Management, Islamabad, Pakistan.

Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic method. Psychological Bulletin, 99, 388-399.

Bennett, N., O'Hare, E., & Lee, J. (1983). Mixed age classes in primary schools: A survey of practice. British Educational Research Journal, 9, 41-56.

Black, G. (1982). Improving teacher performance in rural schools: The case of Sabah, Malaysia. Education for Development, 7, 3-15.

Brandsma, H. P. (1993). Basisschoolkenmerken en de kwaliteit van het onderwijs

Characteristics of primary schools and the quality of instruction

. Groningen, The Netherlands: RION, Instituut voor Onderwijsonderzoek.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 328-375). New York: Macmillan.

Brown, K. G., & Martin, A. B. (1989). Student achievement in multigrade and single grade classes. Education Canada, 29(2), 10-13, 47.

Burchyett, J. A. (1972). A comparison of the effects of nongraded, multi-age, team teaching vs. the modified self-contained classroom at the elementary school level. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Campbell, G. (1993). Assessing the impact of multigrade classes: An annotated bibliography. Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes, 49, 345-364.

Carleton Board of Education. (1991). Summary of observations of a selected group of exemplary multigrade classes in the C.B.E. Nepean. Ontario, Canada: Author.

Carter, J. B. (1973).A study of the effects of multi-grade grouping on the attitudes toward self, peers, and school of selected third and fifth grade students. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Casparis, C., Bernhard, P., & Heuberger, H. (1981). Die Mehrklassenschule, Stiefkind der Schulpoitik? Erste Ergebnisse einer Befragung schweizerischer Mehrklassenlehrer

Multigraded schools, school policy's step child? First results of a Swiss survey

. Schweizerische Lehrerzeitung, 126, 1038-1042.

Chace, E. S. (1961). An analysis of some aspects of multiple-grade grouping in an elementary school. Unpublished doctoral dissertation, University of Tennessee.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. Review of Educational Research, 64, 1-35.

Cohen, J. (1969). Statistical power analysis for behavioral sciences. New York: Academic Press.

Commissie Evaluatie Basisonderwijs. (1994). Inhoud en opbrengsten van het basisonderwijs

Contents and results of primary education

. Leiden, The Netherlands: Distributiecentrm DOP.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design & analysis for field settings. Chicago: Rand McNally.

Daniel, I. H. (1988). "Doing the splits": Core French in the elementary schools. The impact of combined classes on students. Canadian Modern Language Revue/Revue Canadienne des Langues Vivantes, 45, 146-154.

Dodendorf, D. M. (1983). A unique rural school environment. Psychology in the Schools, 20(1), 99-104.

Draisey, A. G. (1985). Vertical grouping in the primary school: A positive view. Education for Development, 9(1), 3-11.

Dreier, W. H. (1949). The differential achievement in rural graded and ungraded school pupils. Journal of Educational Research, 43, 175-185.

Eames, F. H. (1989). A study of the effectiveness of instruction in multi-age grading vs. traditional single-grade level organization on the reading achievement of fourth graders. Unpublished master's thesis, Western Connecticut State University. (ERIC Document Reproduction Service No. ED 309 388)

Evertson, C. M. (1989). Classroom organization and management. In M. C. Reynolds (Ed.), Knowledge base for the beginning teacher (pp. 59-70). Oxford: Pergamon Press.

Finley, C. J., & Thompson. J. M. (1963). A comparison of the achievement of multi-graded and single-graded rural elementary school children. Journal of Educational Research, 56, 471-75.

Fippinger, F. (1967). Empirische Untersuchung zur Leistung von Schuelern aus voll und wenig gegliederten Schulen

An empirical study of scholastic achievement of pupils in fully and partially grade-differentiated schools

. Schule und Psychologie, 14(4), 97-104.

Ford, B. E. (1977). Multiage grouping in the elementary school and children's affective development: A recent review of research. Elementary School Journal, 78(2), 149-159.

Ford, S. (1980). Meeting students' needs in the multiage group environment (E.S.E.A. Title IV-C final evaluation report). Washington, DC: District of Columbia Public Schools. (ERIC Document Reproduction Service No. ED 184 206)

Ford, S. (1981). Meeting students' needs in the multiage group environment (E.S.E.A. Title IV-C final evaluation report, 1979-1980). Washington, DC: District of Columbia Public Schools. (ERIC Document Reproduction Service No. ED 201 709)

Franklin, M. P. (1967). Multigrading in elementary education. Childhood Education, 43(9), 513-515.

Furch-Krafft. E. (1979). Schulische organisation im Primarbereich und Leistungsverhalten: Ein Vergleich der Schulleistungen von Schuelern aus voll und wenig gegliederten Grundschulen

Primary school organization and achievement behavior: A comparison of scholastic achievement with pupils in fully and partially grade-differentiated primary schools

. Zeitschrift fuer Empirische Paedagogik, 3, 297-308.

Gajadharsingh, J. L. (1987). The cognitive profiles of students in multi-grade and single-grade classrooms as a function of gender and locale. Regina, Saskatchewan, Canada: Saskatchewan School Trustees Association Research Centre.

Galton, M., & Simon, B. (Eds.). (1980). Progress and performance in the primary classroom. London: Routledge & Kegan Paul.

Gayfer, M. (Ed.) (1991). The multi-grade classroom: Myth and reality. A Canadian study. Toronto, Ontario, Canada: Canadian Education Association.

Givens, H., Jr. (1972). A comparative study of achievement and attitudinal characteristics of black and white intermediate pupils in individualized-multigrade and self-contained instructional programs. Unpublished doctoral dissertation, Saint Louis University, Saint Louis, Missouri.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

Grove, E. H. (1978). An investigation of the effect of multi-age grouping on student attitude. Unpublished doctoral dissertation, Washington State University.

Gutierrez, R., & Slavin, R. E. (1992). Achievement effects of the nongraded elementary school: A best evidence synthesis. Review of Educational Research, 62, 333-376.

Hammack, B. A. (1974). Self-concept: Evaluation of preschool children in single and multi-age classroom settings. Unpublished doctoral dissertation, Texas Woman's University, Denton.

Harvey, S. B. (1974). A comparison of kindergarten children in multigrade and traditional settings on self-concept, social-emotional development, readiness development, and achievement. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg.

Her Majesty's Inspectorate. (1978). Primary education in England: A survey by HM Inspectors of Schools. London: HMSO, Department of Education and Science.

Her Majesty's Inspectorate. (1982). The new teacher in school. London: HMSO, Department of Education and Science.

Higgins, J. J. (1980). A comparative study between the reading achievement levels of students in a combination/ungraded class and students in a graded class. Unpublished doctoral dissertation, George Peabody College for Teachers of Vanderbilt University, Nashville, TN.

Hoen, R. R. (1972). An evaluation of multi-age classes at Carnavon School, 1971-72. Vancouver, British Columbia, Canada: Board of School Trustees, Department of Planning and Evaluation. (ERIC Document Reproduction Service No. ED 076 649)

Hoffman, L. R. (1981). Multiage grouping and team-teaching: Implications for adaptive instruction. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.

Hopkins, D., & Ellis, P. D. (1991). The effective small primary school: Some significant factors. School Organisation, 11(1), 115-122.

Inspectie Basisonderwijs. (1978). Onderwijsverslag: Verslag van de staat van het onderwijs in Nederland

Report on the state of Dutch education

. The Hague, The Netharlands: Staatsuitgeverij.

Jarousse, J. P., & Mingat, A. (1991a). Efficacite pedagogique de l'enseignement a cours multiples dans le contexte africain

Pedagogical effectiveness of multigrade classes in the context of Africa

. Dijon, France: Institut de Recherche sur L'Economie de l'Education, Centre National de la Recherche Scientifique.

Jarousse, J. P., & Mingat, A. (1991b). Les facteurs agissant sur les acquisitions des eleves a l'ecole primaire au Burkina Faso

Factors influencing the achievement of primary school students in Burkina Faso

. Dijon, France: Universite de Bourgogne, Institut de Recherche sur L'Economie de L'Education.

Jarousse, J. P., & Mingat, A. (1992). L'Ecole primaire en Afrique: Fonctionnement, qualite, produits. Le cas du Togo

The primary school in Africa: Function, quality, results. The case of Togo

. Dijon, France: Universite de Bourgogne, Institut de Recherche sur L'Economie de L'Education.

Jokinen, H. (1979). Yhdysluokkakoulujen oppilaiden koulusaavutuksista 2: Kouluhallituksen yhteisiin kokeisiin perustuva selvitys

On the achievement of pupils in schools with combined grades: Part 2. A study based on centrally administered achievement tests arranged by the National Board of General Education

. Jyvaeskylae, Finland: University of Jyvaeskylae, Institute for Educational Research.

Junell, J. S. (1971). An analysis of the effects of multigrading on a number of noncognitive variables. Unpublished doctoral dissertation, University of Washington, Seattle.

Katz, L. G., Evangelou, D., Hartman, J. A. (1993). The case for mixed-age grouping in early education (3rd ed.). Washington, DC: National Association for the Education of Young Children.

Knight, E. E. (1938). A study of double grades in New Haven city schools. Journal of Experimental Education, 7(1), 11-18.

Knoerzer, W. (1984). Kombinierte Klassen--mehr als eine Notloesung: Ergebnisse eines Forschungsprojekts

Combination classes--more than a temporary solution: Results of a research project

. Paedagogische Welt, 38(3), 144-151.

Knoerzer, W. (1985). Sind Schueler in kombinierten Grundschulklassen benachteiligt? Eine empirische Untersuchung

Are students in multigraded schools put at a disadvantage? An empirical study

. Baltmannsweiler, Germany: Paedagogischer Verlag.

Kral, M. (1995). Effecten van schoolgrootte: Een onderzoek naar effectiviteitsverschillen tussen combinatieklassen en enkelvoudige klassen in het basisonderwijs

Effects of school size: Effectiveness of combination classes and single-grade classes in primary education

. Nijmegen, The Netherlands: Instituut voor Toegepaste Sociale Wet