11-1-2002

# Differences in Achievement Between Graded and Nongraded Elementary School Students

Amanda M. Arkfeld
*University of Nebraska at Omaha*

# DIFFERENCES IN ACHIEVEMENT BETWEEN GRADED AND NONGRADED ELEMENTARY SCHOOL STUDENTS

An Ed.S. Field Project

Presented to the

Department of Psychology

and the

Faculty of the Graduate College

University of Nebraska

In Partial Fulfillment

of the Requirements for the Degree

Specialist In Education (Ed.S.)

University of Nebraska at Omaha

by

Amanda M. Arkfeld

November 2002





ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

THESIS (OR THESIS-EQUIVALENT PROJECT)
(OR EDS FIELD PROJECT) ACCEPTANCE

Acceptance for the faculty of the Graduate College, University of Nebraska, in partial fulfillment of the requirements for the degree (name of degree), University of Nebraska at Omaha.

Committee

James M. Thomas

Leon D. Dippin

Chairperson Alexa Kelly-Vance

Date 1/25/03

Acknowledgements

I would like to thank my Ed.S. Committee Chair, Dr. Lisa Kelly-Vance, for all of her guidance and insight throughout the completion of this research project. Her feedback was always welcome and timely, and I appreciate all of her assistance. I would also like to thank Dr. Leon Dappen and Dr. Jim Thomas for serving on my committee. Each of them provided me with kind feedback that was justified and intuitive.

In addition to my committee members, I would also like to thank my very wonderful research assistants who helped me collect and score my data. Without Eva Denton, Molly Dotson, Melissa Fredenburg, Korrinda Mendez, Kelly Branecki, and Jolene Johnson, data collection would not have even been possible! I appreciate their assistance more than they will ever know. Also, a special thanks goes out to Nicole Werth, Jess Gregory, Jen Johnson, Janet Miller, Kyle Hesser, Lisa Sharpe, and Korrinda Mendez for their assistance in the grueling job of scoring data. Their quick and unselfish assistance will never be forgotten.

The Council Bluffs Community School District also deserves special recognition, because, without their permission to collect data in their district, this research would have never been feasible. Thank you, also, to Stephanie Schmitz, who used her professional connections in the Council Bluffs Community School District to make the data collection process go that much smoother.

Finally, I would like to thank my family and my husband, Sam, for their continued love and unconditional support. Their confidence in me never wanes, and it is because of them that I have been fortunate enough to accomplish all that I have.

# DIFFERENCES IN ACHIEVEMENT BETWEEN GRADED AND NONGRADED ELEMENTARY SCHOOL STUDENTS

Amanda M. Arkfeld, Ed.S.

University of Nebraska, 2004

Advisor: Lisa Kelly-Vance, Ph.D.

Nongraded schools have recently regained popularity. A variety of researchers have compared student achievement levels between nongraded and traditional graded programs to determine efficacy, however results have been inconsistent. The current researcher further examined achievement differences between nongraded and graded classrooms through replication and extension of a previous study performed by Kelly-Vance, Caster, and Ruane (2000), in order to contribute additional research in an area lacking in consistent support. Students were recruited from two schools (i.e., Schools A and B) containing comparable percentages of students receiving free or reduced-cost lunch. Students from School A served as a graded control group, while students from School B were recruited from graded and nongraded classrooms. Dependent measures consisted of student performance on Curriculum-Based Measurement (CBM) probes in the areas of reading, math, and written language. Overall, nongraded students performed as well as, but not better than, graded students on CBM probes. While significant results indicated that nongraded students outperformed graded students at School B in reading at the fifth-sixth grade level and in math at the third-fourth grade level, graded students at School A performed similarly, suggesting that the performance of the nongraded students

could not generalize to other nearby schools. Implications of the current study are discussed.

Table of Contents

# Differences in Achievement Between Graded and Nongraded Elementary School Students

As the concern over appropriate education for American children continues to rise, several schools across the country have implemented an educational approach that at one time was nationwide: the nongraded classroom. Over a century ago, one-room schools were the norm, and children of various ages were educated together from the beginning of their education to the end. Students were typically grouped based on ability level rather than age, and therefore were often educated with students who were a number of years older or younger than themselves. However, as the 20th century approached, the popularity of graded education in public schools increased, and the one-room schoolhouses became a thing of the past (Yarborough & Johnson, 2000).

It was not until the 1960s that the nongraded phenomenon once again was popularized in American schools (Anderson & Pavan, 1993). European advocates of developmentally-oriented education initiated this trend, and soon America saw a considerable rise of nongradedness in its own schools (Anderson, 1992). A decline in this trend was experienced in the early 1980s, but educational interest has again reappeared in the area of nongraded education in present day schools of the 21st century (Yarborough & Johnson, 2000).

In fact, several states including Kentucky, Oregon, and Mississippi have even mandated nongraded education in their primary schools (Heins, Tichenor, Coggins, & Hutchinson, 2000; Lodish, 1992; Pavan, 1992; Yarborough & Johnson, 2000). Due to the rise in popularity of nongradedness, Mason and Stimson (1996) examined 571

elementary schools across 12 states in order to obtain a more specific estimate of the frequency of nongraded classrooms nationwide. Although their results indicated that nongraded classrooms were used relatively infrequently in the United States (i.e., less than 1% of classrooms investigated), they advocated for more research in the area of nongraded education given the recent legislation mandating these classrooms.

Therefore, the present researcher intended to contribute more research to nongraded education, with a specific investigation of achievement differences between nongraded and traditional graded classrooms. Mason and Stimson (1996) argue that few methodologically sound studies have examined achievement differences between these two types of classrooms, and more research is necessary in order to determine if nongraded classrooms are indeed beneficial. In addition, it is essential for educators and administrators to obtain information about potential differences between graded and nongraded classrooms, in order to provide students with more advantageous educational opportunities.

*Graded and Nongraded Schools*

Graded classrooms are the most common and familiar in education. Students are grouped together based on age, and usually one teacher is responsible for one classroom of students. Typically a student will progress through grades first through twelfth in twelve years, given no retentions or skipped grade levels. The majority of schools in the United States tend to utilize a graded approach to education under the assumption that students of the same age can benefit from a similar education (Anderson & Pavan, 1993).

Nongraded classrooms, on the other hand, have been defined in several different ways, but the defining characteristic is that students are grouped together based on ability level rather than age. Traditional grade levels are not used and groups of students within the same classroom will often vary in age. One nongraded classroom may have two or more teachers working together to teach a range of skill levels, in order to accommodate individual differences among students within that classroom; although the student-teacher ratio remains similar to a traditional graded classroom (Gutierrez & Slavin, 1992; Pavan, 1992). In addition, because nongraded classrooms do not limit students to traditional grade level curriculum standards, students typically progress through the grades at their own pace (Heins et al., 2000; Yarborough & Johnson, 2000). As a result, students in the nongraded classroom are rarely retained a whole grade level at once, and advanced students may progress through the classroom curriculum standards at a rapid and appropriate rate. Although some schools utilize nongradedness throughout every classroom, partially nongraded schools and even nongraded subject areas (e.g., reading) have been reported (Gutierrez & Slavin, 1992).

Research in the area of nongraded education is often mistakenly grouped into the general topical area of multiage education. However, it is important to recognize that nongraded classrooms or schools founded under a philosophical approach of nongraded education are truly distinct from multiage or combination classrooms. Multiage or combination classrooms are similar to nongraded classrooms, in that both types of classrooms group students of two or more age groups together within one classroom. However, the important distinction between them lies in how student ability levels are

accommodated in the classroom. As mentioned earlier, students in nongraded classrooms are grouped together based on ability level rather than age. However, students in multiage or combination classrooms are often educated with peers of similar age, while ability levels are not necessarily considered. Multiage or combination classrooms are often implemented on a necessity or convenience basis, rather than on the same philosophical approach of nongraded education.

*Rationale for Implementing Nongraded Schools*

Goodland and Anderson (1987) provided several dimensions of nongraded schools that serve as a rationale for implementing a nongraded program. First, because nongraded programs group students across varying age levels, cooperative learning is promoted. More advanced students can assist those students who are less advanced, whether older or younger than themselves, in order to encourage a more successful classroom environment. Second, nongradedness follows a developmental approach and accounts for differences in student readiness when acquiring new skills. Graded education assumes that students of the same age also generally possess the same skills, when in fact student ability level can vary greatly among children of the same age (Gaustad, 1992; Petrie & Lindauer, 1998). Nongraded classrooms recognize these differences and teachers encourage students to work at their own pace and ability level in the classroom. Finally, because teachers in the nongraded classroom support and expect individual differences among students, they also attempt to provide each student with individualized learning materials that will maximize his or her potential. Therefore,

nongraded classrooms not only acknowledge individual differences among students, but support and welcome them as well.

Other researchers have also investigated the potential benefits of nongraded schools. For example, Guarino (1982) compared one graded and one nongraded school and matched students for age, gender, and intelligence. Results indicated that students in the nongraded school exhibited higher academic achievement and self-concepts and lower anxiety when compared to students in the graded school. Likewise, Anderson and Pavan (1993) stated that students in nongraded schools tend to score as well as or better on measures of academic achievement and mental health than students in graded classrooms. More recent research has also concluded that nongraded students report higher levels of social skills than their graded counterparts (Kelly-Vance et al., 2000).

Pavan (1992) summarized information obtained from 64 studies comparing students from graded and nongraded schools on various measures including academic achievement and mental health. The combined research indicated that (a) nongraded students performed as well as or better on measures of achievement and mental health; (b) the benefits of nongraded education increase over time; and (c) diverse populations (e.g., low socioeconomic level students, boys, African-American students, and underachievers) benefit from a nongraded program.

*Achievement in Nongraded versus Graded Schools*

Although the overall advantages of nongraded schools are well documented, more specific research on achievement differences between nongraded and graded schools are less consistent. The past few decades have yielded several studies indicating that

achievement in nongraded schools is better or at least equal to that of graded schools (Pavan, 1992; Pratt, 1983). For example, Lawson (1973) investigated differences in reading achievement between first, third, and fifth year students in graded and nongraded classrooms, and found that nongraded students performed significantly better on the California Achievement Test versus their graded counterparts. In addition, when Tanner and Decotis (1995) randomly assigned kindergarten and first grade students to graded and nongraded classrooms, they also found supportive results. While no differences were found between the kindergarteners on the Georgia Kindergarten Assessment Program, first grade students in the nongraded classroom earned significantly better grades than their matched peers in the graded classroom.

On the other hand, Veenman (1995) concluded that multiage groupings, including nongraded classrooms, are relatively equivalent to traditional graded classrooms in terms of academic achievement. In his review of 56 studies, Veenman obtained an effect size of close to zero when comparing these groups, and suggested that multiage groupings may essentially be no worse or no better than the traditional single grade groupings that exist in the majority of schools in this nation.

Matthews, Monsaas, and Penick (1997) found similar results using a pretest-posttest design. They investigated the academic achievement of educationally at-risk kindergarten through second grade students in graded and nongraded classrooms in the subject areas of reading and language development. Dependent measures included (a) Iowa Test of Basic Skills (ITBS) scores; (b) a story retelling procedure; (c) Peabody Picture Vocabulary Test-Revised (PPVT-R) scores; and (d) a T-unit analysis of oral

language. They found no significant differences between graded and nongraded student achievement on any of the measures. However, it was noted that nongraded students initially scored lower on measures of reading literacy, although the analyses would have adjusted for these initial differences between groups, and the different instructional environments may have contained variables that the researchers did not control in the results.

Although several studies have concluded that no achievement differences exist between graded and nongraded classrooms, other researchers have found different results. For example, Gutierrez and Slavin (1992) conducted a best evidence synthesis of available research on the achievement effects of nongraded programs at the elementary level, and found consistent positive achievement effects favoring nongraded over graded classrooms. More specifically, nongraded programs were most successful if ample time of direct instruction from the teacher was provided to students, rather than individualized instruction. Students tended to exhibit higher academic achievement if teachers in the nongraded program utilized direct group instruction paired with aspects of individualized instruction that is customary in most nongraded programs.

In addition, more recent investigations into achievement effects in nongraded classrooms have also provided support for their implementation. Kelly-Vance et al. (2000) examined differences in academic achievement in the areas of reading, math, written language, and spelling between graded and nongraded students in grades 1 through 6. Nongraded students were assessed in the existing levels of primary (grades 1, 2, and 3) and intermediate (grades 4, 5, and 6), and graded students were in traditional

grade-level groupings. Dependent measures included scores on Curriculum Based Measurement (CBM) probes in the areas of (a) reading, (b) math, (c) written language, and (d) spelling. Results indicated that the nongraded primary students obtained higher scores in the areas of reading and math compared to students in grades 1, 2, and 3, although no differences between these groups were found in the areas of written language and spelling. Reverse results were found at the intermediate level. Nongraded intermediate students obtained higher scores in the areas of written language and spelling than students in grades 3, 4, and 5, although no differences were found in the areas of reading and math. Overall, none of the comparisons of academic achievement between the graded and nongraded students found the graded students to be superior, leading to quite contradictory results when compared to other studies claiming that nongraded programs are inferior or equal to graded programs.

In summary, the available research in the area of achievement differences between nongraded and graded classrooms remains complicated. Although much evidence is available to support nongraded programs, discrepant information exists. Therefore, it is essential that more research be conducted in this area in order to determine if nongradedness does indeed result in elevated levels of student achievement.

*Measuring Achievement*

In the vast majority of studies investigating achievement differences between nongraded and graded schools, standardized, norm-referenced achievement tests are typically used to measure learner outcome (Gutierrez & Slavin, 1992; Matthews, et al., 1997; Pavan, 1992). However, because achievement tests were originally designed to

measure individual differences among students, there is actually little indication that these tests can also be used to measure academic progress in a student (Hively & Reynolds, 1975). Consequently, this is problematic in that standardized achievement tests may not be sensitive to student academic gain and progress over time (Carver, 1974). In addition, many achievement tests may not actually measure what is in the student curriculum, which as a result may complicate or confound a student's overall score (Shinn, 1989).

CBM has been advocated as an effective alternative to standardized, norm-referenced measures of achievement. CBM uses actual student curricula to assess learner outcome, and therefore achievement effects may be measured in a much more direct manner (Shinn, 1989). CBM procedures have at least five potential advantages over standardized testing. First, CBM is directly tied to student curriculum, therefore allowing the child to be tested over exactly what he or she has had the opportunity to learn in the classroom. Second, CBM allows for quick administration. Each reading probe, for example, takes only one minute to administer by the educator, which in turn permits frequent monitoring of student performance. Third, multiple forms are available. CBM probes are quick and easy to create since they are formed directly from student curriculum, so teachers can have many forms on hand at a given time. Fourth, CBM materials are inexpensive to produce and use. Finally, CBM measurements are sensitive to student improvement over time allowing for accurate data to be captured in pretest-posttest designs (Shinn, 1989).

The current study utilized CBM procedures as an alternative to standardized, norm-referenced tests to assess student achievement in graded and nongraded classrooms based on the supportive research regarding these methods. In particular, CBM is relevant for the needs of the current study, because (a) CBM probes are tied directly to student curriculum, (b) CBM allows for quick administration so that participating students are not absent from the classroom for an extended period of time, and (c) CBM materials are inexpensive to produce and use.

*Rationale for the Current Study*

Several researchers have advocated for further investigation into the effects of nongraded education (Gutierrez & Slavin, 1992; Kelly-Vance et al., 2000; Wong, Erickson, King, Stoller, & Allen 1977). The available research on achievement effects in nongraded classrooms is clearly incongruent, and further examination of the differences between graded and nongraded classrooms will provide a more concise indication of which is superior in terms of educational benefits for students. Additionally, because many recent articles have provided descriptive rather than empirical support for nongraded education, studies using sound methodology are critically important (Mason & Stimson, 1996).

Therefore, the purpose of the current study was to further investigate differences in achievement between graded and nongraded classrooms in order to contribute additional research in an area that is lacking in consistent support. Yarborough and Johnson (2000) advocated that future research should strive to create a simple, straightforward list of protocols for implementing a nongraded program to assist

educators and administrators. Only more research in this area will bring researchers closer to this insightful goal.

Because replication in experimentation has been a long-standing indication of robust results, the current researcher proposes to replicate and expand upon the previous study performed by Kelly-Vance et al. (2000). In a research area plagued with inconsistency, a replicated study is deemed not only beneficial, but also necessary. In fact, Kalat (1999) suggested that when researchers cannot obtain a dependable result, they tend not to accept that result. Therefore, due to the discrepant nature of the available research on achievement effects in nongraded classrooms, the results of a replicated study may assist researchers and educators in determining if nongraded education is indeed effective.

Although the main purpose of the current study was to replicate the previous study performed by Kelly-Vance et al. (2000), two methodological differences existed in the current study that may or may not have affected the results. One difference involved how nongraded students were grouped within the classroom. While Kelly-Vance et al. utilized nongraded students in two existing groups (i.e., grades 1, 2, and 3, and grades 4, 5, and 6), the current study utilized nongraded students in their existing groups of (a) grades 1 and 2, (b) grades 3 and 4, and (c) grades 5 and 6. Although Kelly-Vance et al. and the current researcher both investigated differences in achievement between nongraded and graded classrooms, it is important to recognize that grouping protocols of nongraded students vary across schools.

Mason and Stimson (1996) argue that research regarding organizational structure of nongraded classrooms and descriptions of how students are assigned to and grouped within nongraded classrooms is lacking. Although previous researchers have examined achievement effects between ability groupings in the nongraded classroom (Slavin, 1987), research pertaining to actual classroom assignment is unavailable. Therefore, while this methodological difference existed between the current study and the study performed by Kelly-Vance et al. (2000), this difference may or may not have affected the results.

The second methodological difference that existed between the current study and the Kelly-Vance et al. (2000) study concerned the use of comparison groups. Kelly-Vance et al. utilized a total of two groups to compare achievement differences between graded and nongraded programs: Students from a graded school (School A) and students from a comparable nongraded school (School B). However, in the current study, a total of three groups were utilized: (a) students from a graded school (School A), (b) students from nongraded classrooms in School B, and (c) students from graded classrooms in School B. Therefore, in the current study, achievement differences were not only assessed between graded (School A) and nongraded classrooms (School B) between schools, but also between graded and nongraded classrooms within the same school (School B). The addition of another graded comparison group was expected to strengthen the results of the current study by ruling out any unique differences that may exist between Schools A and B.

*Research Questions and Hypotheses*

The following research question guided the current study: How does the achievement of students in nongraded groups of (a) grades 1 and 2, (b) grades 3 and 4, and (c) grades 5 and 6, differ from the achievement of students in regular graded classrooms in the areas of reading, math, and written language?

Because the purpose of the current study was to replicate the previous study performed by Kelly-Vance et al. (2000), it was hypothesized that the current study would yield similar results, including higher student achievement in nongraded versus graded classrooms overall (i.e., from Schools A and B). With the exception of the two methodological differences mentioned above, a special effort was made by the researcher to replicate all possible aspects of the Kelly-Vance et al. study in order to obtain similar results in the current study. Based on results obtained by Kelly-Vance et al., the current researcher hypothesized that nongraded students in grades 1 and 2 would perform better in the areas of reading and math than students in graded classrooms, due to the performance exhibited by primary level students (grades 1, 2, and 3) in the Kelly-Vance et al. study. In addition, because the intermediate level students (grades 4, 5, and 6) in Kelly-Vance et al. study performed better in written expression than graded students, similar results were expected for nongraded students in grades 5 and 6 in the current study.

Hypothesized results pertaining to nongraded student performance in grades 3 and 4 were less informed. Kelly-Vance et al. (2000) utilized existing nongraded student groups that separated students in the ability levels of grades 3 and 4, while the current

researcher utilized nongraded groups where they were paired together. Although research is lacking in the area of achievement differences between nongraded groups (Mason & Stimson, 1996), it was hypothesized that nongraded students in grades 3 and 4 would more closely resemble student achievement in grades 5 and 6, as opposed to student achievement in grades 1 and 2. Kelly-Vance et al. found that nongraded students in grades 5 and 6 performed better in written expression than their graded counterparts. Because previous research has indicated that written language curriculum in grades 1 and 2 aims to set the foundation for writing skills in later grades (Knudson, 1995), it is likely that nongraded students in grades 3 and 4 would more closely resemble nongraded students in grades 5 and 6 in terms of written expression performance. Furthermore, Knudson (1995) suggested that a benchmark for students in grades 3 and 4 is reaching automatization in writing.

Nongraded students participating in the current study were compared to two distinct groups of graded students on the measure of academic achievement. Nongraded students not only were compared to their graded counterparts attending the same school (School B), but also to other graded students attending a nearby elementary school (School A) that had similar student demographics. The current researcher chose this method to ensure the generalizability of the potential results of the study beyond the uniqueness of School B.

## Method

### *Participants*

Following parental consent and student assent, 182 students from two elementary schools in the same Midwest district participated in this study. The racial composition of the district was comprised of 93% Caucasian, 3% Hispanic, 2% African American, 1% American Indian or Alaskan Native, and 1% Asian or Pacific Islander. Students from School A (containing only graded classrooms) were sampled from grades one through six. Students from School B (containing graded and nongraded classrooms) were sampled from both nongraded and graded first through sixth grade classrooms. See Tables 1, 2, and 3 for demographic information by grade level.

### *Setting*

The two schools chosen for this study were deemed comparable due to the similarity of students receiving free and reduced-cost lunch in each of the schools (School A: 45%; School B: 51%). Both schools also had similar support staff including a school psychologist, a speech-language pathologist, and a Chapter I teacher for specialized reading purposes. The special education (15%) and non-English speaking student populations (0%) were identical at both schools, and each practiced under an inclusion model of special education. Schools A and B also resided in a lower socioeconomic section of the city, and were within a three-mile radius from one another. In addition, School A and School B regularly used curriculum-based measurement (CBM) to monitor regular and special education student progress according to local district CBM norms.

Because Schools A and B resided in the same school district, the same curriculum was used at both schools for the subject areas of reading, math, and written language. The principal at each school reported that the main instructional approach in the graded classrooms was direct instruction, which involved a lecture-format given by the teacher at the front of the classroom. The nongraded classrooms in School B grouped students by ability level, and therefore multiple small groups existed in the classroom. Consequently, the nongraded teacher(s) often provided small group instruction to students, rather than direct instruction to the class as a whole.

School A contained students in kindergarten through sixth grade, including a partial day preschool classroom for students who were considered at-risk for future concerns. The school utilized traditional graded-level groupings, and had a total of 432 students enrolled at the time of the study. School A had 16 regular classrooms and employed 17 teachers. A traditional approach to teaching and curriculum was utilized, with large classroom instruction as the norm.

School B also contained graded kindergarten through grade six classrooms with a total of 482 students enrolled, and was the only school within its district with an all day preschool program for at-risk students. School B also contained nongraded classrooms, which was the focus of the current study. While School B contained 13 regular graded classrooms, 6 nongraded classrooms were also housed within the school. The breakdown of the nongraded classrooms of School B was as follows: (a) one first-second grade classroom, (b) two third-fourth grade classrooms, and (c) three fifth-sixth grade classrooms.

The nongraded program at School B had been implemented for approximately ten years, and was employed under a nongraded philosophical basis. While initially the entire school was nongraded, only recently had the majority of the classrooms returned to a traditional graded model, due to a change in administrative support for nongraded classrooms in the district. Of the existing six nongraded classrooms, team teaching and informal cooperative learning continued to be fundamental to the nongraded model of instruction. In addition, peer helpers and individualized instruction were emphasized within each nongraded classroom. Curriculum materials also had an individualized emphasis to support differentiated student skill levels. No specific training was required of teachers in the nongraded classrooms, and teachers were assigned to the nongraded classrooms primarily on a volunteer basis. No guidelines existed in School B that divided students between the available graded and nongraded classrooms.

*Materials*

Curriculum-based measurement (CBM) was used to measure individual student achievement in reading, math, and written expression at both schools. Because measurement experts continue to challenge the adequacy of standardized measures of achievement in measuring learner outcome, CBM was used as an effective alternative. When compared to standardized achievement measures, CBM has been demonstrated as a more direct way to measure a student's specific skill level in the areas of reading, math, and written expression (Shinn, 1989).

Because of its efficiency in measurement, the reliability of CBM has been well documented in the research. Shinn (1989) summarized various studies that utilized CBM

as a measure of achievement in order to provide a synthesis of its reliability. Test-retest reliability coefficients for each of the subject areas of interest were as follows: (a) reading = .82 to .97; (b) math = .78 to .93; and (c) written expression measured in total words written = .42 to .91. Alternate-form reliability for reading (.84 to .96); math (.48 to .72); and written expression: total words written (.42 to .95) were also provided. Inter-rater reliability coefficients were high within each subject area (reading .99, math .90 to .99, and written expression .98).

The validity of CBM has also been documented in the literature. The construct validity of CBM is most frequently measured by comparing its ability to measure achievement with standardized tests designed for the same purpose. While numerous studies have investigated the construct validity of CBM, generally correlation coefficients for reading ranged from .60 to .80. The median correlation coefficient for math was .43, while the coefficient for written expression ranged from .41 to .84 (Shinn, 1989). Although the correlation coefficients for math and written expression were lower than reading, it has been proposed that standardized tests may be a poor criterion measure in these subject areas (Freeman, Kuhs, Porter, Floden, Schmidt, & Schwille, 1983).

*Procedure*

Following parental consent and student assent, students from Schools A and B were administered CBM probes in reading, math, and written expression. Approximately 10 students from each graded classroom and 20 students from each nongraded classroom were selected. This sampling method was determined based on previous CBM research indicating that a sample of 10 students from a given classroom generally provides a

sufficient estimation of the achievement level for the class at large (Tindal, Germann, & Deno, 1983). Participating students were administered math and written expression probes in small groups with other participating classmates, and reading probes were administered on an individual basis to each student. The primary researcher and all research assistants were trained on administration procedures as proposed by Shinn (1989). The procedures were administered in a standardized method for each student through use of an instruction script (see Appendix), and all students were assessed during school hours.

The district's current CBM probes in reading, math, and written expression were used in this study. The probes have been used in previous research (Kelly-Vance, et al., 2000) and were developed by and administered in the district using the Shinn (1989) method. Each participating student received a CBM packet containing probes in the subject areas of reading, math, and written expression, along with a demographics questionnaire to obtain general information about each student (e.g., age, grade, and gender). Probes were administered in the order of (a) math, (b) written expression, and (c) reading for each participant. Although counterbalancing was considered, it was deemed unnecessary because of lack of effect in previous research (Kelly-Vance, et al., 2000). Each probe was scored using Shinn (1989) scoring procedures. The administration procedures of probes in each subject area are described below.

Reading probes were administered individually to each student. The student was asked to read three separate loose-leaf reading passages during three separate one-minute trials, while the administrator timed the student and marked any reading errors in the

student's CBM packet. Any omitted or incorrectly pronounced words were included in the student score. The administrator computed the amount of words read correctly per minute for each reading passage, and therefore obtained three individual scores for each student. The median of each student's trio of scores was then selected as the student's most representable reading fluency.

Unlike reading, math probes were administered in a group setting. Math probes were developed using grade-level curriculum for grades one through six. Grades 1, 2, and 3 were provided with probes in addition, subtraction, and mixed (a combination of addition and subtraction) problems. Grades 4, 5, and 6 were provided with probes in multiplication, division, and mixed (a combination of addition, subtraction, multiplication, and division) problems. Each student was administered the three grade-appropriate math probes and was given two minutes to complete as many items as possible on each probe. Student responses were scored by counting the correct number of digits provided by the student for each problem. For example, if the correct answer to a problem was 3678, a student received credit for three digits if she provided the answer of 4678.

Written expression probes were also administered in a group setting. Students were asked to listen to a story starter about a grade-appropriate topic. For example, first grade students could have been given the story starter of "Everyone left for lunch and the chalk began to talk." Students were asked to first think about the topic for one minute, and then were given three minutes to write about the topic. The total number of words

written was computed for each student. Words did not need to be spelled correctly to be included in the score.

*Inter-rater Reliability*

Approximately 10% of CBM packets from each grade level were randomly chosen to be scored for inter-rater reliability purposes. A trained research assistant scored the CBM packets following the initial scoring by the primary researcher and other research assistants. Overall, inter-rater reliability was calculated to be .92 across the three subject areas of reading, math, and written language. The following formula was used for inter-rater reliability calculation: number of agreements / number of agreements + number of disagreements × 100. An exact number match between the initial score and the reliability score constituted an agreement (e.g., 96 and 96). However, any initial score and reliability score that was not an exact match was counted as a disagreement (e.g., 96 and 95). When a disagreement occurred, the CBM packet was again checked to determine whether the initial score or the reliability score would be used in data analysis.

In the area of reading, 5 out of 57 possible scores were counted as disagreements (e.g., .92 reliability). In the area of math, 4 out of 57 possible scores were counted as disagreements (e.g., .93 reliability). In the area of written expression, 2 out of 19 possible scores were counted as disagreements (e.g., .89 reliability). The large majority of disagreements in each of the areas involved a difference of two or fewer digits between the initial score and the reliability score. For example, in reading, a disagreement occurred when the reliability research assistant obtained a score of 77 words per minute and the researcher obtained a score of 76 on the same reading probe. The impact of such

a disagreement is minimal, in that only one word out of 77 possible words on the reading probe resulted in the entire probe being labeled as a disagreement.

*Data Analysis*

Data were analyzed using an SPSS statistical package. A series of One-Way Analyses of Variance (ANOVA) were used, along with the computation of descriptive statistics. The dependent variables in this study were student overall scores in the subject areas of reading, math, and written expression. Score differences between graded and nongraded students at Schools A and B were assessed by combining grade levels into three groups: (a) first-second grade students, (b) third-fourth grade students, and (c) fifth-sixth grade students. Analyses for each group were conducted separately. An LSD post-hoc test was used to assess the magnitude of significant results. An alpha level of .05 was considered significant for all analyses.

## Results

A summary of the results is illustrated in Table 1. Means and standard deviations for the first-second, third-fourth, and fifth-sixth grade groups are presented separately in Tables 2, 3 and 4, respectively. The results of graded and nongraded student performance in each CBM subject area are presented below.

*Reading.* At the first-second grade level, no differences were found in the area of reading between graded students from Schools A and B and nongraded students from School B, $F$=.296, $p$=.745. Likewise, no overall differences were found between graded students from Schools A and B and nongraded students from School B at the third-fourth grade level, $F$=3.063, $p$=.055. At the fifth-sixth grade level, differences in reading were

found between graded students from School A and B and nongraded students from School B, $F$=4.346, $p$=.017. An LSD post hoc test revealed that graded students from School A and nongraded students from School B obtained higher scores overall in reading than fifth-sixth graded students from School B ($p < .02$).

*Math.* At the first-second grade level, no overall differences were found in the area of math between graded students from Schools A and B and nongraded students from School B, $F$=.2.757, $p$=.073. At the third-fourth grade level, differences in math were found between graded students from Schools A and B and nongraded students from School B, $F$=6.048, $p$=.004. An LSD post hoc test revealed that graded students from School A and nongraded students from School B obtained higher scores overall in math than graded students from School B ($p < .02$). No overall differences were found between graded students from Schools A and B and nongraded students from School B at the fifth-sixth grade level, $F$=2.938, $p$=.060.

*Written Expression.* No analyses for written expression were found to be significant. Results of the first-second ($F$=.453, $p$=.638), third-fourth ($F$=.216, $p$=.807), and fifth-sixth grade ($F$=1.161, $p$=.319) analyses indicated that no differences existed between graded students from Schools A and B and nongraded students from School B.

## Discussion

Due to the lack of methodologically sound studies in the area of nongraded education, Mason and Stimson (1996) argued that the addition of empirically based studies is necessary for educators to determine if nongraded classrooms are indeed beneficial. The current researcher provided additional research in the area of nongraded

education to assist in the determination of its efficacy, and provided a unique, empirically-based methodology in the utilization of a graded control group. Overall, results indicated that nongraded students performed as well as, but not better than, graded students across Schools A and B. A discussion of the hypotheses and the implications of the current study are presented below.

Because the purpose of the current study was to replicate and extend the findings of Kelly-Vance et al. (2000), it was hypothesized that results in the current study would resemble results found previously by Kelly-Vance et al. First, it was hypothesized that nongraded students would perform better than graded students in reading and math in grades 1 and 2, but have a similar performance in written expression. Second, nongraded students in grades 3 through 6 would perform better in written expression, but have a similar performance in reading and math.

Overall, the results of the current study suggest that neither hypothesis was supported. When comparing graded and nongraded students in grades 1 and 2, no differences were found between the students in the areas of reading, math, and written language across Schools A and B. In grades 3 and 4, no differences were found between graded and nongraded students in the areas of reading and written expression. In the area of math, nongraded students from School B and graded students from School A outperformed graded students from School B. However, because nongraded students did not outperform *all* graded students, including control students from School A, the current researcher suggests there are no overall differences between graded and nongraded students at this grade level. In grades 5 and 6, no differences were found between graded

and nongraded students in the area of math and written expression, but nongraded students from School B and graded students from School A did outperform graded students from School A in the area of reading. However, again, because nongraded students did not exhibit superior performance to *all* graded students, no overall differences are suggested at this grade level in all three of the academic areas assessed.

Noteworthy here, however, is the fact that none of the comparisons between graded and nongraded students suggest that graded students exhibited higher academic achievement than nongraded students. Likewise, while Kelly-Vance et al. (2000) found that nongraded students outperformed graded students in some CBM areas, overall nongraded students were also never outperformed by graded students. Therefore, both studies provide evidence that nongraded students perform at least as well as their graded counterparts in the area of academic achievement.

Several previous researchers have found similar results supporting the notion that nongraded students perform as well as graded students, but not necessarily better. After reviewing 56 studies, Veenman (1995) concluded that multiage, including nongraded, groupings may essentially be no worse or no better than traditional grade level groupings. Matthews, et al. (1997) also found no differences between graded and nongraded kindergarten through second grade students on a number of dependent measures. In addition, when comparing 57 previous studies in the area of nongraded education, Pavan (1992) found that nearly one-third of the studies suggested that nongraded students performed only at a comparable, but not superior, level to graded students. Therefore, perhaps the reason behind the multitude of incongruent studies either promoting or

opposing nongradedness in the area of academic achievement, is due to the fact that nongraded education does not consistently increase achievement outcomes in students overall compared to graded education.

The importance of the use of a control group (School A) in the current study should be highlighted for this reason. Had there been no control group, nongraded students from School B would have appeared academically superior to their graded counterparts in School B in the areas of math in grades 3 and 4 and in reading in grades 5 and 6. The control students from School A served an important function illustrating that nongraded students from School B only outperformed graded students within the unique environment of School B. In addition, the control group demonstrated that the results could not be generalized to another nearby school within the same school district that was matched on the variables of ethnicity and socioeconomic status. Since many previous studies in the area of nongraded education failed to use a control group when favoring academic performance in the nongraded classroom, perhaps future researchers should strive to replicate findings with the addition of a control group in their methodology.

Future researchers examining the area of nongraded education are also advised to investigate the role of the teacher in influencing student academic achievement. Perhaps nongraded educators possess a higher level of enthusiasm for teaching compared to traditional graded educators, and this enthusiasm translates into higher academic motivation in nongraded students. The nongraded educator may indeed be masking, or at least contributing to, the true effect of nongraded academic achievement. In addition, because nongraded students often remain with the same teacher for multiple school years

compared to students in the traditional graded model of education, the number of years a student is taught by the same teacher should also be examined. Finally, future researchers may also wish to examine gender differences in nongraded education, since previous researchers have documented significant differences between male and female student achievement (Kelly-Vance, et al., 2000; Pavan, 1992). The current researcher was unable to perform gender analyses in the current study due to the lack of a sufficient number of participants.

Educators and administrators who are considering the implementation of nongraded education in their schools can benefit from the results of the current study in the following ways. First, it is suggested that academic achievement will likely be comparable between graded and nongraded elementary school classrooms. Therefore, school districts may implement the program that best fits their needs. For example, previous research has documented other social benefits of nongradedness (Kelly-Vance, et al., 2000; Pavan, 1992), and more emphasis should perhaps be placed on these benefits rather than academic achievement. Second, it is evident that the specific way in which nongraded students are grouped is not critical. Previous researchers have utilized and assessed a variety of existing nongraded groups in student achievement, and yet have obtained similar results that correspond with the current study.

In conclusion, the current study supports the notion that nongraded students perform as well as their graded counterparts on measures of academic achievement, but do not outperform them. Previous research in the area of academic achievement in graded versus nongraded education has been inconsistent, and perhaps this is due to the fact that

no true differences exist at all. Only further research in this area, utilizing sound methodology and the use control groups, will help clarify this confounded area of research.

References

Anderson, R. H. (1992). *The nongraded elementary school: Lessons from history.* (ERIC Document Reproduction Service No. ED 348161)

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

Carver, R. P. (1974). Two dimensions of tests: Psychometric and edumetric. *American Psychologist, 29,* 512-518.

Freeman, D. J., Kuhs, T. M., Porter, A. C., Floden, R. E., Schmidt, W. H., & Schwille, J. R. (1983). Do textbooks and tests define a national curriculum in elementary school mathematics? *Elementary School Journal, 83,* 501-513.

Gaustad, J. (1992). *Making the transition from graded to nongraded primary education.* OSSC Bulletin, 35(8) entire issue.

Goodland, J. I., & Anderson, R. H. (1987). *The nongraded elementary school* (rev. ed.). New York: Teachers College Press.

Guarino, A. R. (1982). *An investigation of achievement, self-concept, and school related anxiety in graded and nongraded elementary schools.* Unpublished Doctoral Dissertation, Rutgers 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.

Heins, E. D., Tichenor, M. S., Coggins, C. J., & Hutchinson, C. J. (2000). Multiage classrooms: Putting theory into practice. *Contemporary Education, 71,* 30-36.

Hively, W., & Reynolds, M. C. (Eds.). (1975). *Domain reference testing in special education*. Reston, VA: The Council for Exceptional Children.

Kalat, J. W. (Ed.). (1999). *Introduction to Psychology*. Belmont: Wadsworth.

Kelly-Vance, L., Caster, A., & Ruane, A. (2000). Nongraded versus graded elementary schools: An analysis of achievement and social skills. *The Alberta Journal of Educational Research, 46,* 372-390.

Knudson, R. E. (1995). Writing experiences, attitudes, and achievement of first to sixth graders. *The Journal of Educational Research, 89,* 90-97.

Lawson, R. E. (1973). *A comparison of the development of self-concept and achievement in reading of students in the first, third, and fifth year of attendance in graded and nongraded elementary schools*. Unpublished doctoral dissertation, Ball State University, Muncie, IN.

Lodish, R. (1992). The pros and cons of mixed-age grouping. *Principal, 71,* 20-22.

Mason, D. A., & Stimson, J. (1996). Combination and nongraded classes: Definitions and frequency in twelve states. *The Elementary School Journal, 96,* 439-452.

Matthews, M. W., Monsaas, J. A., & Penick, J. M. (1997). A comparative study of literacy development of at-risk children in graded versus nongraded classrooms. *Reading Research and Instruction, 36,* 225-239.

Pavan, B. N. (1992). The benefits of nongraded schools. *Educational Leadership, 50,* 22-25.

Petrie, G., & Lindauer, P. (1998). Parents, community, and the nongraded primary: Making a successful transition. *Contemporary Education, 69,* 92-94.

Pratt, D. (1983). *Age Segregation in Schools.* Montreal, Canada: Paper presented at the annual meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED 231038)

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children.* New York: Guilford Press.

Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best evidence synthesis. *Review of Educational Research, 57,* 293-336.

Tanner, C. K., & Decotis, J. D. (1995). The effects of continuous-progress nongraded primary school programs on student performance and attitudes toward learning. *Journal of Research and Development in Education, 28,* 135-143.

Tindal, G., Germann, G., & Deno, S. L. (1983). *Descriptive research on the Pine County norms: A compilation of findings* (Research Report No. 132). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.

Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: A best evidence synthesis. *Review of Educational Research, 65,* 319-381.

Wong, M. R., Erickson, B., King, P., Stoller, J., & Allen T. (1977). Perceptions and activities in an open education environment. *Child Study Journal, 7,* 117-130.

Yarborough, B. H., & Johnson, R. A. (2000). Nongraded schools: Why their promise has not been realized and should be reconsidered. *Contemporary Education, 71,* 42-48.

# Appendix

*Instruction Script for Reading Probes*

Say to the student: "When I say 'start,' begin reading aloud at the top of this page. Read across the page [demonstrate by pointing]. Try to read each word. If you come to a word you don't know, I'll tell it to you. Be sure to do your best reading. Are there any questions?"

Say "Start."

After 1 minute, say "Stop."

*Instruction Script for Math Probes*

Say to the student: "The sheets on your desk are math facts."

For single-skill probes say: "All of the problems are [addition or subtraction or multiplication or division] facts."

For multiple-skill probes say: "There are several types of problems on the sheet. Some are addition, some are subtraction, some are multiplication, and some are division [as appropriate]. Look at each problem carefully before you answer it."

"When I say 'start,' Turn them over and begin answering the problems. Start on the first problem on the left on the top row [point]. Work across and then go to the next row. If you can't answer the problem make an 'X' on it and go to the next one. If you finish one page, go to the next one. Are there any questions?"

Say "Start."

After 2 minutes, say "Stop."

*Instruction Script for Written Expression Probes*

Say to the student: "I want you to write a story. I am going to read a sentence to you first, and then I want you to write a short story about what happens. You will have 1 minute to think about the story you will write and then have 3 minutes to write it. Do your best work. If you don't know how to spell a word, you should guess. Are there any questions?"

"For the next minute, think about…[insert the story starter]."

After 1 minute is up, say "Start writing."

After 3 minutes, say "Stop and put your pencil down."

(Shinn, 1989)

Table 1

*Summary of ANOVA Results Between Schools A and B by Grade level*

| Grade | Reading | | Math | | Written Language | |
|---|---|---|---|---|---|---|
| | *F* | *p* | *F* | *p* | *F* | *p* |
| Grades 1 and 2 | .296 | .745 | 2.757 | .073 | .453 | .638 |
| Grades 3 and 4 | 3.063 | .055 | 6.048 | .004** | .216 | .807 |
| Grades 5 and 6 | 4.346 | .017* | 2.938 | .060 | 1.161 | .319 |

*Note.* *$p < .05$. **$p < .01$.

Table 2

*Means, Standard Deviations, and Demographics for First-Second Grade Students by School*

| | School A | School B | |
|---|---|---|---|
| | Graded | Nongraded | Graded |
| Reading | 73.82 (43.03) | 80.06 (44.09) | 83.88 (37.66) |
| Math | 96.86 (45.67) | 66.66 (25.57) | 87.23 (38.43) |
| Written Expression | 24.47 (11.38) | 24.06 (8.32) | 21.64 (8.14) |
| n | 23 | 15 | 17 |
| Males | 14 | 4 | 7 |
| Females | 9 | 11 | 10 |

Table 3

*Means, Standard Deviations, and Demographics for Third-Fourth Grade Students by School*

| | School A | School B | |
|---|---|---|---|
| | Graded | Nongraded | Graded |
| Reading | 109.37 (25.70) | 105.86 (43.15) | 80.15 (43.78) |
| Math | 143.25 (48.26) | 150.59 (41.49) | 102.42 (50.56) |
| Written Expression | 40.18 (14.57) | 41.54 (12.52) | 38.31 (19.53) |
| n | 16 | 22 | 19 |
| Males | 7 | 6 | 4 |
| Females | 9 | 16 | 15 |

Table 4

*Means, Standard Deviations, and Demographics for Fifth-Sixth Grade Students by School*

| | School A | School B | |
|---|---|---|---|
| | Graded | Nongraded | Graded |
| Reading | 125.80 (40.62) | 127.52 (36.58) | 96.16 (44.89) |
| Math | 177.44 (80.06) | 154.90 (53.79) | 129.41 (69.30) |
| Written Expression | 46.52 (13.36) | 51.23 (9.64) | 45.37 (16.33) |
| n | 25 | 21 | 24 |
| Males | 13 | 10 | 9 |
| Females | 12 | 11 | 15 |