


Winter 2019

## The Impact of Grade Level Flexible Grouping on Math Achievement Scores

Heather McKeen

Columbus State University, [mckeen\\_heather@columbusstate.edu](mailto:mckeen_heather@columbusstate.edu)

Follow this and additional works at: <https://digitalcommons.georgiasouthern.edu/gerjournal>

 Part of the [Curriculum and Instruction Commons](#), [Curriculum and Social Inquiry Commons](#), [Early Childhood Education Commons](#), [Educational Methods Commons](#), [Elementary Education Commons](#), and the [Gender Equity in Education Commons](#)

---

### Recommended Citation

McKeen, Heather (2019) "The Impact of Grade Level Flexible Grouping on Math Achievement Scores," *Georgia Educational Researcher*: Vol. 16 : Iss. 1 , Article 5.

Available at: <https://digitalcommons.georgiasouthern.edu/gerjournal/vol16/iss1/5>

This elementary education is brought to you for free and open access by the Journals at Digital Commons@Georgia Southern. It has been accepted for inclusion in Georgia Educational Researcher by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact [digitalcommons@georgiasouthern.edu](mailto:digitalcommons@georgiasouthern.edu).

---

# The Impact of Grade Level Flexible Grouping on Math Achievement Scores

## **Abstract**

Flexible grouping aims to divide students into groups according to their strengths and abilities. This will allow teachers to tailor instruction to meet the needs of students, allowing them to thrive within their current environments. Quantitative research was used to investigate whether flexible grouping had a positive effect on student achievement as measured by the mathematics section of the Criterion-Referenced Competency Test (CRCT). The aggregate scores of students in grades 1 to 4 in a school in a county in Georgia were compared for periods before and after the implementation of flexible grouping. The following study was able to conclude that flexible grouping helps teachers closely monitor students which, in turn, allows lessons to be more catered to individual strengths and weaknesses; however, there was no direct correlation between flexible grouping and performance. There were students who performed well and others who did not. It is therefore recommended that further quantitative research based on survey and experimental designs be conducted at several other schools to corroborate or refute the results of this study for the new Georgia Milestones Mathematics Achievement Scores.

## **Keywords**

Flexible Grouping. “Systematic assessment and on-going observation to formulate students into groups according to specific goals, activities, and individual needs” (Catherine Valentino, 2000). 5 Differentiated Instruction. “A teaching theory based on the premise that instructional approaches should vary and be adapted in relation to individual and diverse students in classrooms” (Tomlinson, 2001). Criterion Referenced Competency Test. “The CRCT is designed to measure how well students acquire, learn, and accomplish the knowledge and skills set forth in a specific curriculum or unit of instruction (Georgia Department of Education, 2008).

## **Creative Commons License**



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

## **Introduction**

Flexible grouping incorporates a variety of learning strategies, including learning through interactions with peers and guidance from superiors. Flexible grouping methods also include both teacher-led and student-led grouping. In teacher-led grouping, students are stratified by the teacher based on ability, interest, or level of skill or content mastery. In this case, there will be more direct guidance because someone with experience will be monitoring the learning. In student grouping, students take control of the grouping process, which involves sub-categorization into collaborative groups, performance-based groups, and student pairs (Conklin, 2007). Students are able to work more independently in this case.

While some research has been conducted on the effectiveness of flexible grouping in achieving positive academic outcomes (Tieso, 2005), it is unclear what effect this approach has had on student achievement in mathematics. Although not widely utilized, flexible grouping has been found to be one of best strategies applicable in classrooms of students with special needs (Fisher, 2011; Subban, 2006). Flexible grouping can be utilized to help students who need more support during the instructional process, or who come to school lacking motivation or interest (“Harris County,” 2012). Teachers have faced the challenge of applying appropriate teaching techniques to improve students’ grades in mathematics and other science-related subjects. One common and preferred method based on flexible grouping entails either placing students in teacher-guided groups or allowing students to lead themselves. In teacher-based grouping, the teacher responds to the ability of the students, while in student-based 10 groupings, students group themselves according to their interests (Teno, 2000).

The following study will utilize this hypothesis in order to discover whether flexible grouping is successful in improving the performances of students in the classrooms. Through an observational approach and analysis of results, such as test scores, this information will be able to help discover whether or not this method of teaching is effective and should be used to supplement other learning techniques. The instruments of measurement for the analysis will be explained in detail in the following sections.

### **Significance of the Study**

Flexible grouping is a relatively new concept that allows a teacher to focus on students who share similar qualities, in an effort to improve the efficiency and effectiveness of instruction. This teaching strategy ensures that students with similar abilities and capabilities are grouped together. Through the use of flexible grouping, both the teacher and the students take control of the learning process,

making it easier to improve student performance. In order to measure student academic performance, CRCT is used; it also determines whether a student is promoted to the next grade level or retained at the same academic level. More specifically, CRCT stands for “Criterion-Referenced Competency Test” and is a system that is implemented into schools to measure how successful a student performs. This study sought to ask whether flexible grouping was connected to student achievement in mathematics, as measured by the CRCT.

As schools and educational decision-makers actively seek effective and efficient strategies to enhance student learning, this study sought to clarify the role that flexible grouping can play in increasing student achievement. Educators are increasingly required to utilize scientifically based instructional practices, so empirical research on three effective strategies is essential. Valentino (2000) and Marzano, Pickering and Pollack (2011) consider flexible grouping to be a suitable method for preparing students for the CRCT, but also an efficient way of improving overall student performance. This study could aid in understanding the gap that exists between the theoretical framework of the strategy and classroom results, and may assist educational decision-making in other schools.

Teachers in the one-room schoolhouse utilized instructional strategies that allowed each learner to study tasks appropriate to his or her level of development. Additionally, more advanced students would teach less proficient students as a way of meeting all the students’ emotional and instructional needs and coping with the inevitable diversity in students’ achievement levels (George, 2005). In this case, it is often very beneficial for students to help each other because they are able to collaborate and learn together. Teachers enhanced the instructional process by encouraging students’ habits of responsibility for their own learning and willingness to help one another learn. Teachers also initiated instructional strategies and routines to maximize cooperation in order that students could be independent and efficient whether in learning individually or collectively (Daniel, 1999).

The Industrial Era brought about new challenges for the American education system. With the urbanization caused by industrialization, new educational attitudes and policies emerged in the U.S. to produce people with the expertise to work in various industries (Laprade, 2010). Economic development resulting from industrialization helped to expand the role and mission of the educational system (George, 2005). As income increased and the economy became more complex, society started placing a higher value on schooling. As books became more widely available, more Americans had access to information, which in turn led to societal transformation (Daniel, 1999). In order to provide a large enough number of highly qualified professionals to support burgeoning industry, a

large segment of the population had to be educated (Valentino, 2000). School leaders acknowledged the power and efficiency of factories and the virtues of a disciplined and orderly taskforce (Daniel, 1999). In America, teachers inspired students to follow directions and submit to authority, as these were primary qualifications needed for workers in mass-production (Daniel, 1999). The U.S. education system thus became rooted in the Protestant work ethic, in which students who worked hard and desisted from misbehavior received a reward (Deniz & Tortora, 2005).

During the 1960s, the civil rights movement had a profound impact on American public education. In an effort to improve the academic performance of all students, schools and teachers began to be held accountable for high academic standards (Laprade, 2010). It was also suggested that minority students did not have the same types of educational opportunities as their white peers, resulting in significant academic achievement gaps. As a result, school segregation was abolished, resulting in a movement towards more diverse and integrated classrooms. However, simply desegregating classrooms did not necessarily improve student achievement. Rather, it led to increased tracking of students according to ability, socio-economic status, race (Daniel, 1999), or the content presented to students (George, 2005).

This history is important to note because it demonstrates how the American school system has evolved over time. Now, classrooms are becoming more flexible and are able to adhere to a more versatile student body in order to work with a variety of different strengths and weaknesses. One of the biggest challenges faced in desegregated and non-tracked classrooms has always been effectively meeting the needs of diverse student populations. When a teacher is faced with this situation, he or she must be prepared to help students thrive no matter their level of expertise. The heterogeneously grouped classroom may show a wide range of student diversity in ability, socio-economic status, cultural/linguistic diversity, learning styles, or previous academic experiences. A lack of linguistic or social integration often leads to lower student achievement (Ozturk & Debelak, 2005). Wilson (2012) asserts that teachers should prepare themselves and students for the challenge of interacting and communicating with different races.

### **Research Questions**

This study will explore and aim to answer the following questions:

How does flexible grouping affect the learning curve of students in the classroom?  
How did flexible grouping directly impact the mathematical scores of the students using the CRCT scale?

What are the differences in effectiveness between peer to peer learning and teacher to peer learning?

### **Conceptual Framework**

Flexible grouping is one strategy being utilized by many schools in Georgia to improve the quality of instruction and learning. Working collaboratively, students can accomplish their tasks by learning from each other (Teno, 2000). In addition, flexible grouping has found a wide range of applications in education, regardless of students' grade level. It has enriched the knowledge students acquire from their teachers and from fellow students who have a better grasp of specific content or skills (Meijnen & Guldemond, 2002).

This study will utilize the Criterion-Referenced Competency Test (CRCT) because it is a uniform basis of measurement. In the past, it has been implemented into course curriculums in order to effectively measure the performance of students and whether or not they should be permitted to pass on to the next grade level. Therefore, the CRCT provides existing data regarding the students in the classrooms being examined. Once flexible grouping has been implemented and performance is measured by the CRCT, there will be a proper basis for comparison in order to note whether the students improved over time. Then, this can help decide whether there is a direct correlation between flexible grouping and improved academic performance.

### **Limitations of the Study**

Educational systems that track students also have some limitations. Students from lower tracks may not be exposed to higher-quality work or advanced material, limiting the models they have to achieve at higher levels. Tracking can also result in stigmatization of low-tracked students, which can affect attitude as well as more tangible aspects of academic performance (Daniel, 1999). Tracking also generally does not allow for students to move into higher academic levels, even if the student has demonstrated ability and interest in any other track (Deniz & Tortora, 2005). The widespread use of tracking started to die out around the 1980's following the realization that "it was giving students in low-track classes less resources, fewer experienced teachers, low expectations, and unchallenging curricula" (Valentino, 2000). Increasingly educators thought that poorly performing students would benefit from sharing a class with better-performing students (Valentino, 2000). Heterogeneous classrooms allow students to engage in a curriculum which allows peer learning and collaboration. Students not only get opportunities to contribute during class sessions and to appreciate classmates' contributions, but according to Valentino (2000), heterogeneously grouped students in foreign countries significantly outperform high-tracked American students. However, heterogeneous

grouping can be unfair to high achievers as they become bored with a lack of challenging assignments and instruction that moves more slowly than their own rate of progress, while low achievers become disadvantaged because failure to keep up with peers may induce low self-esteem (Conklin, 2007).

## **Methods**

The purpose of this study was to apply a causal-comparative design to explore the extent to which flexible grouping affects student achievement in an elementary school. The purposive sample consisted of two pre-existing cohorts of students across grades one through four attending one school in the Harris County School District. The secondary data consisted of the mathematics scores achieved in the end-of-year Criterion Referenced Competency Test (CRCT). The CRCT mathematics scores for the previous cohort, who were never exposed to flexible grouping in the 1st through 4th grades during the academic years 2003-4 to 2007-8, were compared with the CRCT mathematics scores of the current cohort, who were exposed to flexible grouping in the 1st through 4th grades in the academic years 2008-9 to 2011-12. The effects of gender and ethnicity on CRCT mathematics scores were also explored. This chapter justifies the research design, defines the research questions, hypotheses, and variables, and describes the procedures used for data collection and analysis.

Two of the major limitations of this study were that the CRCT may not be an accurate indicator of student performance and the study did not examine the extent to which flexible grouping was implemented in each classroom. The study results may not be generalizable outside of the school system where the study was conducted both because the study was conducted in a single county where only grade levels 1 to 4 were under study and because the research depended heavily on data from secondary sources, which may lack external and internal validity. Also, although every attempt was made to ensure that the results of ANOVA conducted in this study were valid and did not violate theoretical assumptions, the application of inferential statistics to analyze data collected in educational settings to support school management and policy decisions is controversial. Some authors (e.g., Carver, 1993; Daniel, 1998; Schmidt, 1996) have argued that the use of statistical tests for such purposes should be banned. A survey of American Educational Research Association (AERA) members indicated that 19% agreed (Mittag & Thompson, 2000).

There are several reasons to support the argument that null hypothesis testing is flawed and has limited applications in educational research. Due to the severe limitations imposed by the use of a causal-comparative quantitative design to test hypotheses at one school as discussed above, it is essential to conduct further

research to corroborate the results of this study. It is recommended that the same methods that were used in this study, based on secondary data in school archives, should be repeated at several other schools in other districts. If the results of this study can be replicated at several other schools, this would provide more convincing evidence to generalize the finding that flexible grouping is significantly more effective in the 1st through 3rd grades than in the 4th grade, and that the effects of flexible grouping do not vary significantly with respect to student gender and ethnicity. In addition to corroborating the results of this study using a causal-comparative design, it is recommended that more powerful experimental research designs should be implemented. Experimental designs are essential in education to draw conclusions about effects of interventions which cannot be so easily evaluated using non-experimental causal-comparative designs (Fraenkel & Wallen, 2010). An experimental design is imperative to study causal relationships because the researcher can manipulate the causes to generate corresponding effects, in order to answer the research question "What is the effect of flexible grouping on the test scores of students?" An experiment with two randomly selected and assigned groups of students should be performed to determine the extent to which the hypothetical cause (flexible grouping) influences the hypothetical affect (the test scores of the students). Random selection and assignment are necessary in a true experimental study to ensure that the students represent the essential characteristics of the populations from which they were drawn, in terms of their demographic and academic characteristics.

### **Rigor of Methods**

The causal-comparative design applied in this study is defined as "research to explore the cause for, or consequences of existing differences in groups of individuals, also referred to as *ex post facto* research" (Fraenkel & Wallen, 2010). The essential features of this design are that the groups of participants were pre-existing, so they could not be manipulated by the researcher. Because the alleged causes or consequences of flexible grouping have already occurred, and were studied in retrospect, this design is also called *ex post facto* (Latin for "after the fact"). A causal-comparative design is not experimental, because the researcher did not create differences between the groups by manipulating the dependent and independent variables. Nor did the researcher randomly select or assign the participants into groups. The dependent and independent variables were fixed by circumstances that were out of the control of the researcher.

Qualitative research methodologies, underpinned by the social constructivist paradigm, are also recommended to study the effects of flexible grouping. The decision to apply a quantitative research methodology in this study was underpinned by the 67-positivist paradigm, meaning that facts and feelings are



separate, and that academic achievement can be measured objectively, predicted by hypotheses, and summarized by statistics (Fraenkel & Wallen, 2010). Although a quantitative research methodology enables the achievements of mutually exclusive groups of students to be compared in terms of statistics, it cannot explain the multitude of subtleties and nuances that differentiate the achievements of each individual student at a personal level. Because inferential statistics are based on mean values, then all that can be concluded is that, on average, the effect was different among the group of students exposed to flexible grouping, relative to the group who were not. "On average" implies that a substantial proportion (but not all) of the students may be influenced by flexible grouping.

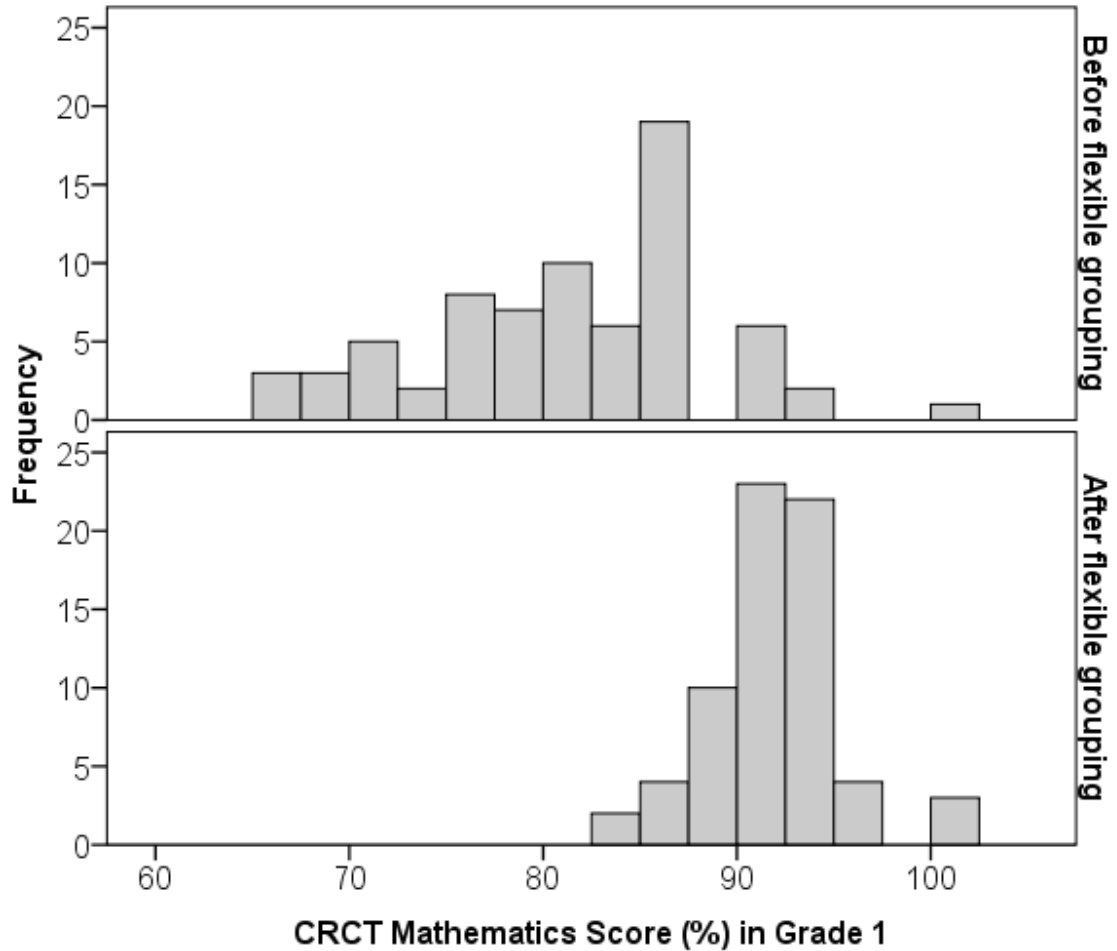
To address this difficulty, the social constructivist paradigm (that facts and feelings are not separate) must be applied. This paradigm assumes that academic performance cannot simply be summarized objectively in terms of statistics, but must be considered in terms of multiple subjective realities, constructed from the many different choices, attitudes, behaviors, and cognitive abilities of each individual student and teacher (Fraenkel & Wallen, 2010; Palinscar, 1998). For example, Erwin (1991) proposed that a broader range of assessment tools is required to evaluate student academic performance and that quantitative analysis of test scores alone does not necessarily provide an accurate indicator. Biggs (1999) and Shepherd (2000) also advocated moving away from teacher-oriented quantitative assessment models to student oriented qualitative models of assessment. Fraenkel and Wallen (2010) emphasized that the main strength of the qualitative analysis of empirical data collected in natural educational settings is that it provides a much more comprehensive view of the performance of teachers and students than do quantitative forms of educational research.

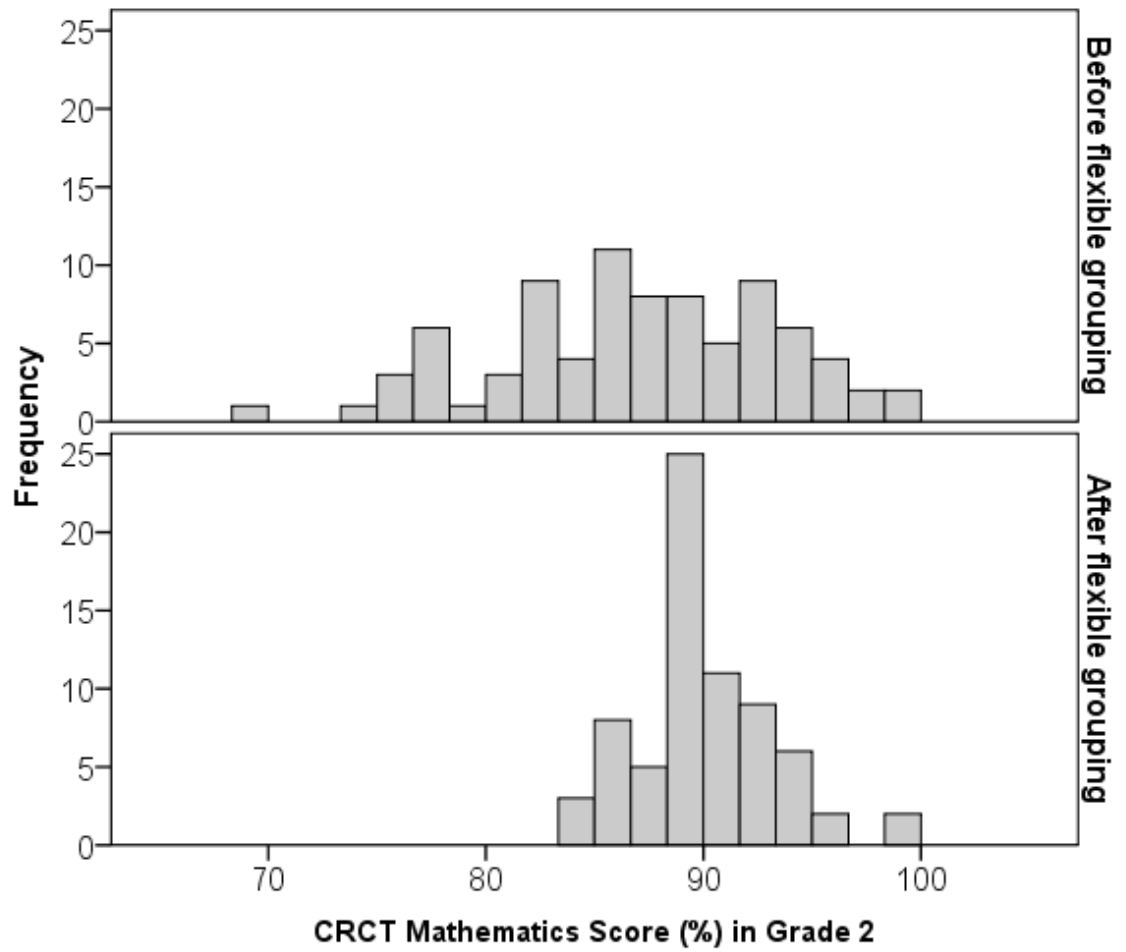
### **Discussion of Findings**

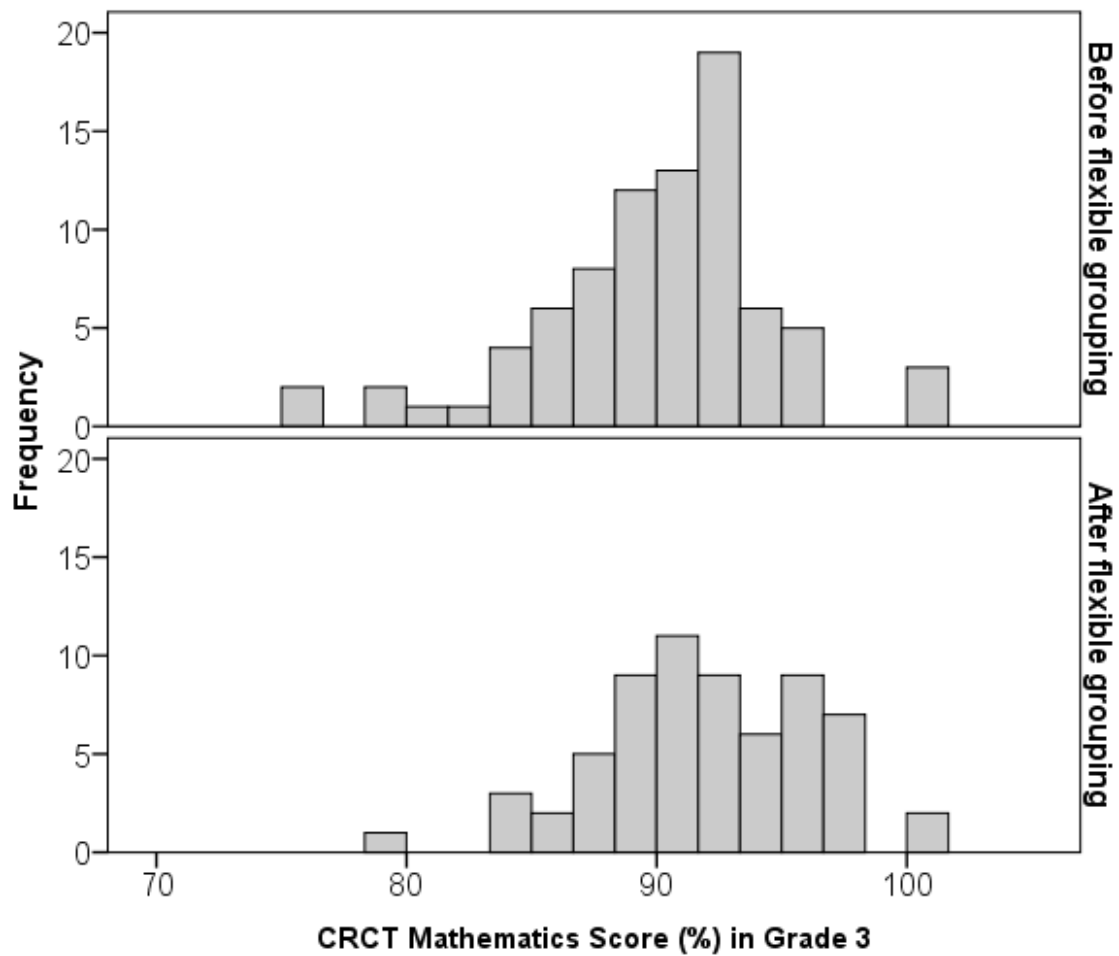
Statistical evidence indicated that the impact of flexible grouping on the CRCT scores in mathematics varied with respect to the grades of the students. The mean CRCT scores of the cohort exposed to flexible grouping were elevated and the variance was reduced relative to the cohort that was not exposed to flexible grouping in the 1st through to the 3rd grades. In the 4th grade, however, no significant effects were found in the CRCT scores of cohorts who were exposed to flexible grouping relative to the cohort that was not so exposed.

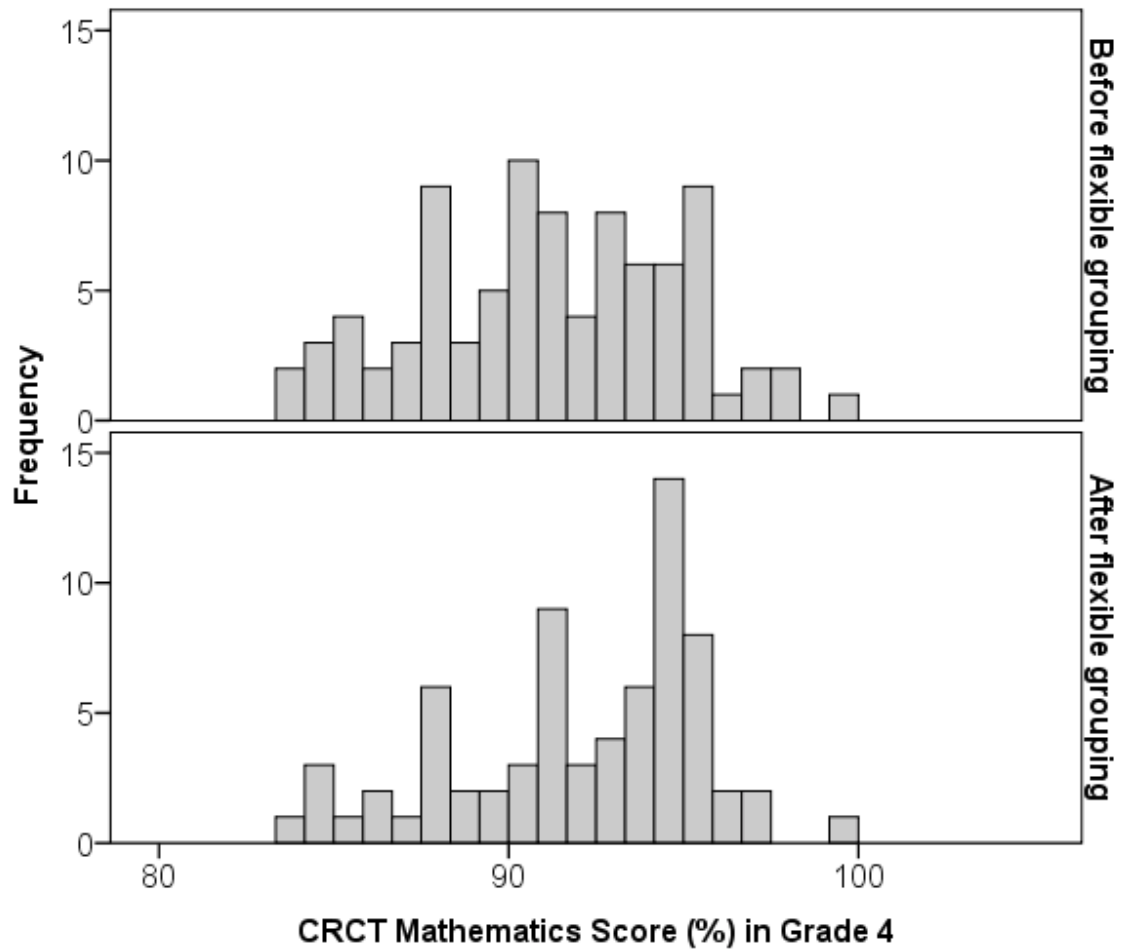
By enabling the matching of ability levels with a skill, flexible grouping provides greater flexibility to meet individual needs (Weaver, 2006). Flexible grouping strategies give teachers time to get to know their students well, provide them with stimulating learning experiences, and help them explore aspects of the world other than those prescribed by the curriculum (Weaver, 2006). If

implemented well, flexible grouping may unleash the greater potential of children learning in the classroom. Using such a strategy, the teacher will at one point instruct the students about the content of their lessons, but not everything about the lessons hinges on the capacity of the teacher to relay information. Importantly, the capacity and interest of the student to learn is viewed as another factor in effective teaching (Weaver, 2006).









**Contribution to Education**

Flexible grouping can also positively influence classroom management, for which the relationship between a teacher and a student is of paramount importance (Stephen, 2011). In a large, heterogeneously-grouped classroom, it may be difficult for some teachers to develop rapport with each student. Flexible grouping strategies can allow teachers continually to assess the student’s performance in and outside the classroom and to spend more directed time with students in small groups; this supports the development of a unique relationship with each student (Heacox, 2002), in turn supporting effective and efficient management of the classroom. Because flexible grouping allows for small-group instruction, it also lets the teacher easily monitor and adjust learning and redirect potential student misbehavior before it becomes problematic (Heinemann & Dunlap, 2005). It ensures the participation of all students because the teacher closely monitors work and provides feedback.

More research needs to be done to determine if the findings will be the same with the GA Milestones as opposed to the CRCT.

**Keywords**

Flexible Grouping. “Systematic assessment and on-going observation to formulate students into groups according to specific goals, activities, and individual needs” (Catherine Valentino, 2000).

Differentiated Instruction. “A teaching theory based on the premise that instructional approaches should vary and be adapted in relation to individual and diverse students in classrooms” (Tomlinson, 2001).

Criterion Referenced Competency Test. “The CRCT is designed to measure how well students acquire, learn, and accomplish the knowledge and skills set forth in a specific curriculum or unit of instruction (Georgia Department of Education, 2008).

## References

- Bender, W. N., & Shores, C. (2007) *Response to intervention: A practical guide for every teacher*. Thousand Oaks, CA: Corwin Press, Inc.
- Biggs, J. (1999). *Assessing for learning quality*. Buckingham UK: Open University.
- Carver, R. (1993). The case against statistical significance testing revisited. *Journal of Experimental Education*, 61, 287-292.
- Christopher, W., & Park, B. (2006). Considering the Tower of Babel: Correlates of assimilation and multiculturalism among ethnic minority and majority groups in the United States. *Social Justice Research*, 19(3), 277-306. doi:10.1007/s11211-006-0014-8
- Conklin, W. (2007). *Applying differentiation strategies: Teacher's handbook for grade 3-5*. Huntington Beach, CA: Shell Education.
- Daniel, L. G. (1998). Statistical significance testing: A historical overview of misuse and misinterpretation with implications for the editorial policies of educational journals. *Research in the Schools*, 5, 23-32
- Deniz, C., & Tortora, M. (2005). Flexible grouping and student learning in a high-needs school. *Education & Urban Society*, 37(2), 139-150. doi:10.1177/0013124504270787
- Erwin, T.D. (1991). *Assessing student learning and development*. San Francisco: Jossey-Bass.
- Fisher, E. (2011). *Effectively teaching in the mainstream classroom: assistive technologies for students with auditory disabilities*. Retrieved from <http://orange.eserver.org/issues/7-3/fischer.html>
- Fraenkel, J. R., & Wallen, N. E. (2010). *How to design and evaluate research in education* (7th ed.). New York, NY: McGraw-Hill.
- George, P. (2005). A rationale for differentiating instruction in the regular classroom. *Theory into Practice*, 44(3), 185-197. DOI:10.1207/s15430421tip4403\_2

- Georgia Department of Education (2008). Georgia Department of Education—  
Assessment: Criterion-referenced competency tests. Atlanta, GA: Author.
- Harris County School District (2012). Retrieved from <http://www.education.com/schoolfinder/us/georgia/district/harris-county-school-district/>
- Heacox, D. (2002). *Differentiating instruction in the regular classroom: How to reach and teach all learners, grades 3-12*. Minneapolis, MN: Free Spirit Publishing, Inc.
- Heinemann, M., & Dunlap, G. (2005). Positive support strategies for students with behavioural disorders in general education settings. *Psychology in the Schools*, 42(8), 779-794. DOI: 10.1002/pits.20112
- Kracl, C. (2011). Managing small group instruction through the implementation of literacy work stations. Retrieved from <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1093&context=cehsdis>
- Laprade, K. (2010). Removing instructional barriers: one track at a time. *Education Journal*, 131(4), 740-752.
- Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Meijnen, G., & Guldmond, H. (2002). Grouping in primary schools and reference processes. *Educational Research & Evaluation*, 8(3), 229-249. doi:10.1076/edre.8.3.229.3857
- Mittag, K. C., & Thompson, B. (2000). A national survey of AERA members perceptions of statistical significance tests and other statistical issues. *Educational Researcher*, 29, 14-20.
- Ozturk, M. A., & Debelak, C. (2005). Setting realistically high academic standards and expectations. Retrieved from <http://www.usca.edu/essays/vol152005/ozturkrev.Pdf>
- Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological*



Methods, 1, 115–129.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29, 4-14.

Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal*, 7(7), 935-947.

Teno, K. (2000). Cluster grouping elementary gifted students in the regular classroom. *Gifted Child Today*, 23(1), 44-51.

Tieso, C. (2005). The effects of grouping practices and curricular adjustments on achievement. *Journal for the Education of the Gifted*, 29(1), 60-89.

Tomlinson, C. A., (2001). *How to differentiate instruction in mixed-ability classrooms*. (2nd Ed.) Alexandria, VA: ASCD.

Valentino, C. (2000). *Flexible grouping*. Washington, DC: National Academy Press.

Weaver, M. (2006). Exploring conceptions of learning and teaching through the creation of flexible learning spaces: The Learning Gateway-A case study. *New Review of Academic Librarianship*, 12(2), 108-125. DOI: 10.1080/13614530701330414

Wilson, K. (2012). Critical multicultural pavilion - research room. Retrieved from <http://www.edchange.org/multicultural/papers/keith.html>