Abstract Title Page

Title: Evaluating the Efficacy of Using Pre-differentiated and Enriched Mathematics Curricula for Grade 3 Students

Authors and Affiliations:

D. Betsy McCoach

E. Jean Gubbins

Jennifer Foreman

University of Connecticut

Karen E. Rambo

Colorado State University

Lisa DaVia Rubenstein

Ball State University

Abstract Body

Background / Context:

Although research on the effectiveness of differentiated and enriched instruction in improving the achievement of diverse students is still emerging, some studies (Beecher & Sweeney, 2008; Brimijoin, 2001; Gavin & Casa, 2012; Tieso, 2002; Tomlinson, Brimijoin, & Narvaez, 2008) suggest that students in academically diverse classrooms benefited academically from differentiated learning experiences. Brighton, Hertberg, Moon, Tomlinson, and Callahan (2005) found modest improvements in all content areas for middle school students involved in differentiated instruction and assessment. Recently, Reis, McCoach, Little, Muller, and Kaniskan (2011) found improvement in reading for one suburban district and oral reading fluency for one suburban school. Additionally, both oral reading fluency and reading comprehension were higher in the treatment group in one low-SES urban school. More research is clearly warranted to assess the effectiveness of differentiated and enriched instruction and enriched curricula.

Purpose / Objective / Research Question / Focus of Study:

The primary research question was "What is the impact of implementing the pre-differentiated mathematics curricula in algebra, geometry and measurement, and graphing and data analysis on the achievement of grade 3 students, after controlling for pretest achievement scores?" Specifically, we were interested in examining whether math achievement outcomes of treatment and control group students differed.

Setting:

The study included 42 public schools, and one private school in 12 states, with the majority from rural setting and 3 schools within a large city. Nine schools had more than 20% non-White/non-Asian student enrollment and 5 schools had more than 30% non-White/non-Asian student enrollment. Free- and reduced-priced meal eligibility for students at these schools ranged from 0 to 68%. In both groups, teachers were predominantly female and White. Both treatment and control teachers had similar characteristics: over 57% had 10 or more years teaching experience; a majority had less than 10 years of experience with grade 3 students; and over 56% had master's degrees.

Population / Participants / Subjects:

The number of treatment and control students in the final analytic sample was 2290. Of the students in the analytic sample, a similar percentage of males (50%) and females (49%) comprised the treatment and control groups across all schools. Over 80% of students in the treatment and control groups were White, with fewer than 20% representing other racial/ethnic groups.

Intervention / Program / Practice:

This study compared researcher-designed, pre-differentiated and enriched mathematics curricula in algebra, geometry, and measurement, graphing, and data analysis to the districts' mathematics curricula.

Three widely adopted models in gifted and talented education place the teacher in the role of knowledge broker, facilitator, and guide, emphasizing differentiation of curricula in general

education classrooms as well as in pull-out and special classes designed for identified gifted and talented students. Elements of these models were combined and utilized to develop the current study's mathematics units: Differentiation of Instruction Model (Tomlinson, 2001); Depth and Complexity Model (Kaplan, 2009); and Schoolwide Enrichment Model (Renzulli & Reis, 1997).

Using pre-assessments accompanying each unit, teachers were guided in their selection of differentiated lesson options, based on the same challenging concepts, appropriate for each student. For most of the units' lessons, three levels of scaffolding were embedded in the lessons' activities. This form of differentiating by students' demonstrated prior knowledge—often known as differentiation by readiness (Tomlinson, 2001, Tomlinson & Jarvis, 2009)—is referred to as "tiering." Tiered activities (Adams & Pierce, 2006; Tomlinson, 2001) function to lead students with different levels of initial knowledge and skills to master a similar "big idea" objective through adjustment of such aspects of the assignment as simplicity/complexity, concreteness/abstractness, more structure/less structure, etc. (Tomlinson, 2001). Treatment teachers participated in 2 days of onsite professional development, completed teacher logs upon completion of each unit, and research team members maintained weekly contact with treatment teachers. Over 90% of treatment teachers completed all three units.

Research Design:

This multisite cluster randomized control trial randomly assigned 141 general education classrooms (teachers) within 43 schools across 12 states to treatment or control conditions. Treatment teachers were required to implement three curricular units, which would supplant the district's adopted mathematics curricula for 16 weeks. Control group teachers continued with the district's adopted mathematics curricula or "business as usual." Of the 141 teachers, 84 were assigned to the treatment condition; 57 assigned to the control condition. In two instances of co-teaching, the co-teachers were assigned to condition as a single unit.

Cluster-level randomization was selected for "good practical and scientific reasons" (Shadish, Cook, & Campbell, 2002, p. 254). On a practical level, participant recruitment required support from school administrators, for which the cluster-level design was pragmatically suited. Scientifically, we hoped to answer questions about real students in real classrooms for whom the layers of clustered data provide nuanced estimates of outcomes.

During the spring prior to the intervention, grade 2 students in the participating schools completed the Level 8 Math Problems subtest or another nationally standardized achievement test. All pretest measures of ability and achievement were aligned using the equipercentile method in which scale scores were converted to z-scores for comparability. Pretest math achievement scores were used as a covariate in the resulting analyses. After the curricular implementation was complete, treatment and control students took the Level 9 ITBS Math Problem Solving and Data Interpretation subtest as a posttest achievement measure—the dependent variable for the 3-level analyses.

Data Collection and Analysis:

Assessments/Measures

Teachers administered one mathematics subtest of the Iowa Tests of Basic Skills (ITBS) to the treatment and control students. The ITBS test content is aligned with the most current content standards, curriculum frameworks, and instructional materials. The test was standardized on a national sample of students K-9, with approximately 3,000 students per level per form completing the tests. Internal consistency estimates using KR 20 varied between .79 and .98.

Students in the standardization sample represented various types of communities, ethnicity, race, and socioeconomic status. The standardization sample included public, parochial, and non-parochial schools. Schools in the standardization were further stratified by socioeconomic status. Data from these sources were used to develop special norms for a variety of groups (e.g., race/ethnicity, public school) (Hoover et al., 2003).

The ITBS Level 8 Math Problems subtest was administered to grade 2 students prior to the curricular intervention to obtain information on students' achievement in mathematics. The Level 8 ITBS subtest had 30 items. A small proportion of students completed other mathematics achievement pretests (the TerraNova, the Measure of Academic Progress [MAP], or the Stanford Achievement Test [SAT]). Because the achievement tests were on different scales, z-scores for the scores on each of the four achievement tests were calculated so that students' pretest achievement could be compared across tests.

Analyses

To examine the effects of the differentiated curricula, we first ran a series of 3-level regression models using HLM 7.0 software (Raudenbush, Bryk, & Congdon, 2011). At level 1, we included pre-ITBS score, which was grand mean centered, and "gifted" status, which was defined as students with CogAT composite IQ scores in the top 10% of their respective schools. At level 2, we included treatment, which was dummy coded, so that 0 represented a control classroom and 1 represented a treatment classroom. At level 3, we controlled for the school mean achievement by creating an aggregate of each school's second grade math pretest score. School aggregate math score was also a z-score. Because the ITBS scores exhibited a ceiling effect, the data were reanalyzed using a multilevel Tobit model, which accounted for the censored nature of the data. The results of the two analyses were quite similar and led to identical conclusions about treatment effectiveness. Table 1 contains the results of the final results from the 3-level multilevel analysis in HLM and the two level multilevel Tobit analyses with corrected standard errors in MPLUS 6. MPLUS 7 now allows for three level organizational analyses, so the data will be rerun using a 3-level Tobit model in MPLUS prior to the presentation, but given the similarity between the current analyses, we do not expect the results to change appreciably.

Findings / Results:

The final model failed to show a main effect for treatment, but did uncover interesting cross-level interaction effects. Examining Model 3, although there was no statistically significant difference between treatment and control groups when school aggregate pre-ITBS was held constant, there was a statistically significant effect of treatment on the pre-ITBS slope, that is, on the effect of pre-ITBS on post-ITBS. The effect of pre-ITBS on post-ITBS was stronger in treatment classes than in control classes, indicating that the treatment appeared to have a differentiating effect on students.

However, the picture is even more complex. The school aggregate pre-ITBS score moderated the cross-level interaction between treatment and pretest score. In schools with lower pre-ITBS scores, the treatment slope was steeper than the control slope; in higher aggregate pre-ITBS schools, this effect was reversed. These 3-way interaction effects are most easily understood graphically. Therefore, Figures 1, 2, and 3 illustrate the relationship between pre-ITBS and post-ITBS scores in low aggregate pre-ITBS schools, high aggregate pre-ITBS schools, and average aggregate pre-ITBS schools. In average aggregate pre-ITBS schools, there appears to be no discernable treatment effect, based on the final HLM models. In low pre-ITBS schools, students with higher pretest scores do better in the treatment group, and students with

lower pretest scores do better in the control group. In high pre-ITBS schools, students with lower pretest scores do better in the treatment group, and students with high pre-ITBS scores appear to do equally well in either group. See Table 1 for the results of the analyses.

To illustrate this in another manner, we broke the group into four groups, based on their relative pretest levels. Table 2 shows the differences between the treatment and control groups disaggregated by their relative standing within their schools, based on their standardized pretest scores. The treatment effect was negligible for average and low achievers. However, there was a difference of .41 standard deviation units, favoring the treatment for the highest achievers. These results suggest that differentiated instruction may be most effective for the highest achievers in a school. This effect was likely strongest for the highest achievers in the lower achieving schools due to the observed ceiling effects on the post ITBS.

Conclusions:

In general, the post-ITBS scores of students in the treatment group were equal to those in the control group. However, high achieving students did appear to derive some benefit from the differentiated curricula. This was especially true for high achieving students in lower achieving schools. Several conclusions can be posited:

1. The ceiling on the norm-referenced test was not high enough to record students' true level of content, concepts, and skills mastered in problem solving and data interpretation.
2. The norm-referenced ITBS was not a good match to content in the algebra and geometry and measurement units.
3. The lack of a main effect illustrated that eliminating 16 weeks of the "business as usual" curricula for the treatment group students did not have a negative impact on students involved in the intervention.
4. The curricula benefited students differentially depending on the achievement status of their schools and their designation as treatment group or control group students.

We were able to replace grade level curriculum with more challenging and enriching curriculum without negatively impacting standardized test scores. In the current age of increased accountability, teachers are often afraid to stray from the mainstream curriculum for fear of jeopardizing their state test scores. Assuming the ITBS posttest measures the typical grade 3 mathematics curricula, the current study provides some evidence that teachers can replace typical at-grade level curriculum with more challenging, enriched mathematics curriculum without suffering adverse consequences on standardized assessments. Viewed through this lens, the results of this study should encourage teachers to consider stepping out of the lock-step curriculum to differentiate their math curriculum.

The measurement issues that plagued this study (i.e.- the low ceiling on the ITBS, the lack of alignment between the ITBS and the differentiated units) are a major limitation. Future research should explore the differentiated units using different post-assessments, including the researcher developed curriculum based measures for both the treatment and control groups and utilizing and out of grade level assessments would have provided a clearer picture of the effects of the intervention.

Appendix A: **References**

Adams, C. & Pierce, R. (2006). *Differentiating instruction: A practical guide to tiered lessons in the elementary grades.* Waco, TX: Prufrock Press.

Assouline, S., & Lupkowski-Shoplik, A. (2005). *Developing math talent: A guide for educating gifted and advanced learners in math.* Waco, TX: Prufrock Press.

Ball, D. L., Hill, H. C, & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator, 29*(1), 14-17, 20-22, 43-46.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). *Content knowledge for teaching: What makes it special? Journal of Teacher Education, 59*(5), 389-407.

Beecher, M., & Sweeny, S. (2008). Closing the achievement gap with curriculum enrichment and differentiation: One school's story. *Journal of Advanced Academics, 19*, 502-530.

Bellanca, J., & Brandt, R. (Eds.). (2010). *21st century skills: Rethinking how students learn.* Bloomington, IN: Solution Tree Press.

Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook 1: Cognitive domain*. New York, NY: Longman Publishers.

Brandwein, P. F. (1983). *Notes toward a renewal in the teaching of science.* New York, NY: Coronado Publishers.

Brighton, C., Hertberg, H., Callahan, C., Tomlinson, C., & Moon, T. (2005). *The feasibility of high end learning in academically diverse middle schools* (RM05210). Storrs, CT: The National Research Center on the Gifted and Talented, University of Connecticut.

Brimijoin, K. (2001). *Expertise in differentiation: A preservice and inservice teacher make their way.* (Unpublished doctoral dissertation). University of Virginia: Charlottesville.

Chapin, S. H., O'Connor, C., & Anderson, N. C. (2003). *Classroom discussions: Using math talk to help students learn, grades 1-6.* Sausalito, CA: Math Solutions

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York, NY: Academic Press.

Davis, E. A., Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning, *Educational Researcher, 34*(3), 3-14.

Farkas, S., & Duffett, A. (2008). *High achieving students in the era of No Child Left Behind.* Washington, DC: The Thomas B. Fordham Institute.

Farkas, S., & Duffett, A. (2010). *Cracks in the ivory tower? The view of education professors circa 2010*. Washington, DC: The Thomas B. Fordham Institute.

Friesen, J. (2003). Giving students 21st century skills: A practical guide to contemporary literacy. *Multimedia Schools, 10*(3), 22-26.

Gavin, M. K., & Casa, T. M. (2012). Nurturing young student mathematicians. *Gifted Education International,* pp. 1-14. doi: 10.1177/0261429412447711

Gavin, M. K., Casa, T. M., Adelson, J. L., Carroll, S. R., & Jensen Sheffield, L. (2009). The impact of advanced curriculum on the achievement of mathematically promising elementary students. *Gifted Child Quarterly, 53,* 188-202.

Haycock, K. (2001). Help all students achieve: Closing the achievement gap. *Educational Leadership, 58*(6), 6-11.

Hertberg-Davis, H. (2009). Myth 7: Differentiation in the regular classroom is equivalent to gifted programs and is sufficient: Classroom teachers have the time, the skill, and the will to differentiate adequately. *Gifted Child Quarterly, 53,* 251-253.

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30*, 393-425.
James, W. (1885). On the functions of cognition. *Mind, 10*, 27-44.
Hoover, H. D., Dunbar, S. B., Frisbie, D. A., Oberley, K. R., Ordman, V. L. Naylor, R. J., . . . Shannon, G. P. (2003). *The Iowa Tests guide to research and development.* Itasca, IL: Riverside Publishing.
Kaplan, S. N. (2009). The grid: A model to construct differentiated curriculum for the gifted. In J. S. Renzulli, E. J. Gubbins, K. S. McMillen, R. D. Eckert, & C. A. Little (Eds.), *Systems & models for developing programs for the gifted & talented* (2nd ed., pp. 235-251). Mansfield Center, CT: Creative Learning Press.
Lastica, J. R., & O'Donnell, C. L. (2007, April). *Considering the role of fidelity of implementation in science education research: Fidelity as teacher and student adherence to structure.* Paper presented at the meeting of the American Educational Research Association, Chicago, IL.
McBee. M. (2010). Modeling outcomes with floor or ceiling effects: An introduction to the Tobit Model. *Gifted Child Quarterly, 54*, 314-320.
Raudenbush, S. W., Bryk, A. S., & Congdon Jr., R. T. (2011). *HLM 7.0 for Windows.* Skokie, IL: Scientific Software International.
Regan, B. (2008). Why we need 21st century skills—and how to do it. *Multimedia & Internet@Schools*, *15*(4), 10-13.
Reis, S. M., McCoach, D. B., Little, C. A., Muller, L. M., & Kaniskan, R. B. (2011). The effects of differentiated instruction and enrichment pedagogy on reading achievement in five elementary schools. *American Educational Research Journal, 48*, 462-501.
Renzulli, J. S., & Reis, S. M. (1997). *The Schoolwide Enrichment Model: A how-to guide for educational excellence* (2nd ed.). Mansfield Center, CT: Creative Learning Press.
Saltpeter, J. (2003). 21st century skills: Will our students be prepared? *Technology & Learning*, *24*(3), 17-18.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth.
Sheffield, L. J. (1999). *Developing mathematically promising students*. Reston, VA: National Council of Teachers of Mathematics.
Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.
Stuart, L. (1999). *21st century skills for 21st century jobs.* Washington, DC: U.S. Department of Commerce, U.S. Department of Education, U.S. Department of Labor, National Institute for Literacy, & Small Business Administration. (ERIC Document Reproduction Service No. ED445249)
Tieso, C. (2002). *The effects of grouping and curricular practices on intermediate students' math achievement* (RM02154). Storrs, CT: The National Research Center on the Gifted and Talented, University of Connecticut.
Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
Tomlinson, C. A., Brimijoin, K., & Narvaez, L. (2008). *The differentiated school: Making revolutionary changes in teaching and learning*. Alexandria, VA: ASCD.
Tomlinson, C. A., & Jarvis, J. M. (2009). Differentiation: Making curriculum work for all students through responsive planning and instruction. In J. S. Renzulli, E. J. Gubbins, K.

S. McMillen, R. D. Eckert, & C. A. Little (Eds.), *Systems & models for developing programs for the gifted & talented* (2nd ed., pp. 599-628). Mansfield Center, CT: Creative Learning Press.

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times.* San Francisco, CA: Jossey-Bass.

Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research, 78,* 516-551.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological process.* Cambridge, MA: Harvard University Press.

Appendix B. Tables and Figures

Table 1
Results of Multilevel Analyses of the Treatment Effect Using a Tobit Model

		HLM Model Coefficient (*SE*)	MPLUS Non-censored	MPLUS Censored
Model for Intercept of post test score (β_{00j})				
Intercept				
	Intercept (γ_{000})	204.17***	205.32***	206.09***
		(1.08)	(.92)	(.99)
	Mean school pretest (γ_{001})	2.90	3.30	4.43
		(3.11)	(3.06)	(3.34)
Treatment				
	Intercept (γ_{010})	.81	.51	.87
		(1.10)	(.96)	(1.08)
	Mean school pretest (γ_{011})	1.12	2.44	1.76
		(3.25)	(2.80)	(3.23)
Model for student achievement slope				
Intercept				
	Intercept (γ_{100})	13.46***	13.31	14.33
		(0.64)	(.64)	(.72)
	Mean school pretest (γ_{101})	2.39	2.01	4.01*
		(1.81)	(1.78)	(1.82)
Treatment				
	Intercept (γ_{110})	2.28	2.26**	2.79**
		(0.82)**	(.75)	(.83)
	Mean school pretest (γ_{111})	-6.92**	-6.11**	-6.77**
		(2.36)	(2.16)	(2.29)
Model for gifted effect				
Intercept				
	Intercept (γ_{200})	6.88***	6.67***	9.89***
		(1.27)	(.91)	(1.24)
Variance				
Level 1 (between students)				
	$Var(e_{ijk})$	245.89	242.31	303.04
		(7.84)	(7.56)	(12.62)
Level 2 (between teachers)				
	$Var(r_{0jk})=\tau\beta$	17.07***	28.96	34.93
		(4.88)	(7.23)	(8.83)

		HLM Model Coefficient (*SE*)	MPLUS Non-censored	MPLUS Censored
Level 3 (between schools)				
	Var(u_{00k})	14.85*** (6.07)		
Goodness of fit				
	AIC	17596.6	19073.7	17580.8
	BIC	17665.4	19136.8	17643.8
	Deviance	17572.6	19051.7	17558.8
	Parameters	12	11	11

Table 2

Mean Posttest Achievement of Treatment and Control Students in Four Categories of Pretest Achievement

Experimental Condition	Pretest Achievement: Low Pretest	Low-Average Pretest	High-Average Pretest	High Pretest	Total
Control					
Group *N*	133	301	338	127	899
Mean	182.66	200.4	213.41	222.60	205.80
Standard Deviation	15.99	19.13	19.09	16.46	22.03
Treatment					
Group *N*	168	523	495	205	1391
Mean	181.34	198.66	214.18	228.61	206.51
Standard Deviation	17.37	19.54	19.00	13.69	22.98
Effect Size					
Cohen's *d*	-.08	-.09	.04	.41	.03

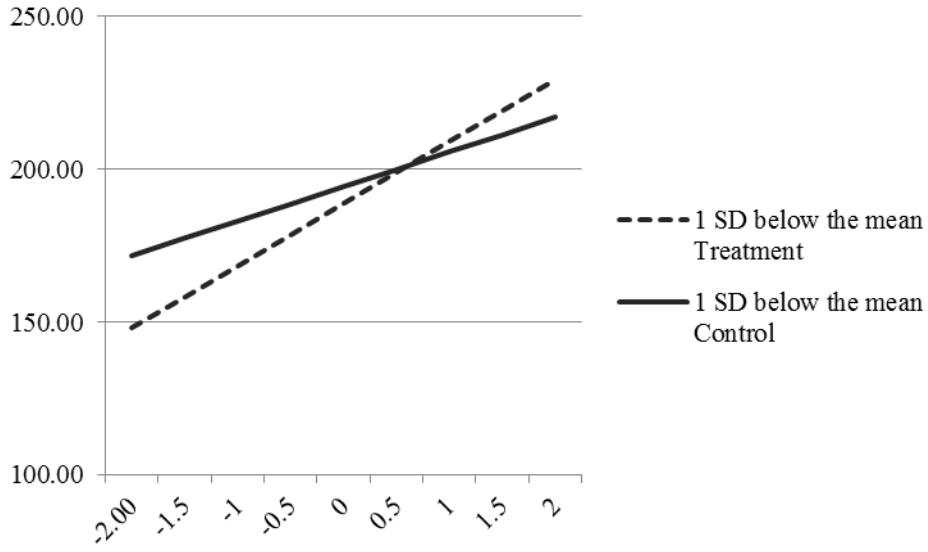

Figure 1. Predicted values for students with a given math pretest score (X-axis) on final math posttest score (Y-axis) in schools that scored one standard deviation below the sample mean on pre math achievement.

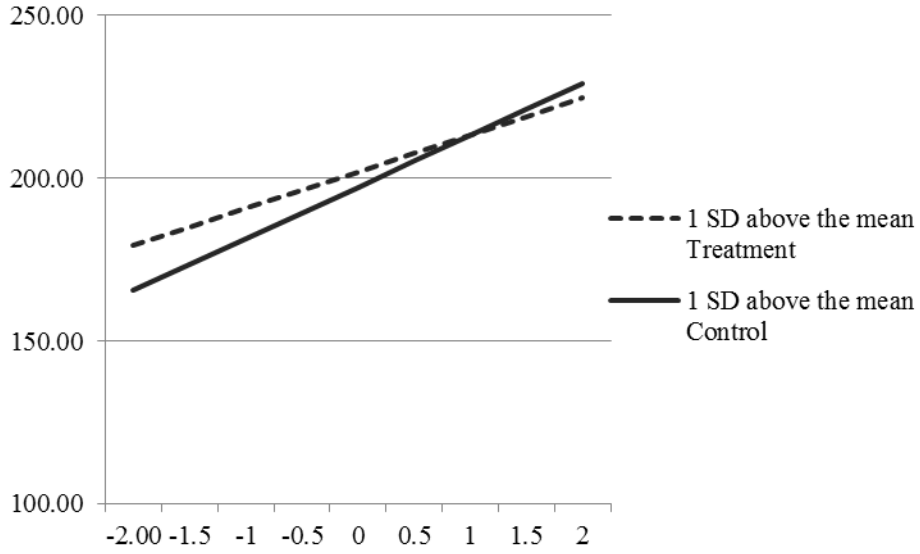

Figure 2. Predicted values for students with a given math pretest score (X-axis) on final math posttest score (Y-axis) in schools that scored one standard deviation above the sample mean on pre math achievement.

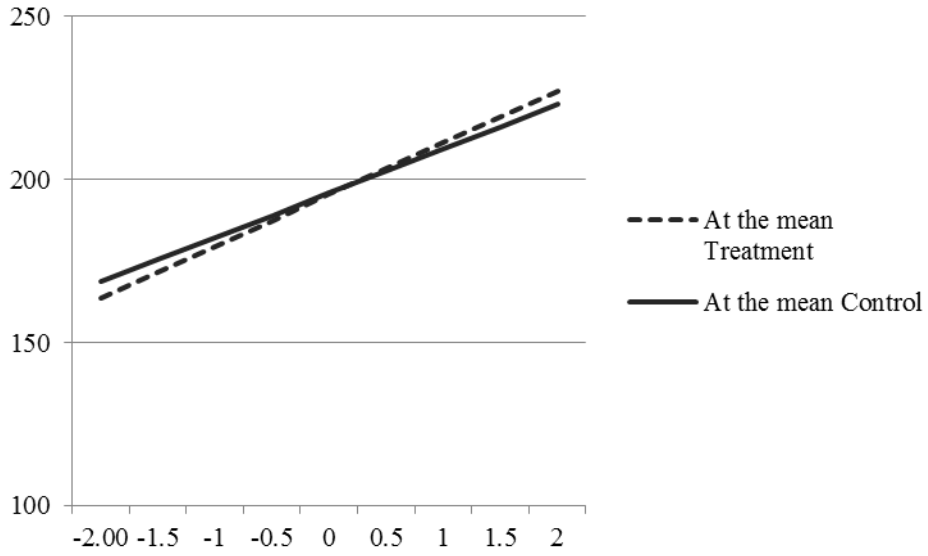

Figure 3. Predicted values for students with a given math pretest score (X-axis) on final math posttest score (Y-axis) in schools that scored at the sample mean on pre math achievement.