Abstract Title Page
Not included in page count.

Title: A Randomized Controlled Trial of Two Online Mathematics Curricula

Author(s): Haiwen Wang, Katrina Woodworth

Abstract Body

Limit 4 pages single spaced.

Background / Context:

Description of prior research and its intellectual context.

This study applies a randomized controlled trial to examine the effects of supplemental instruction using two online mathematics curricula—DreamBox and Reasoning Mind. It is an independent evaluation intended to generate unbiased results that will help inform the ongoing development of a charter school network's hybrid instructional model, which supplements traditional face-to-face instruction with instruction provided via computer-based programs. It also provides evidence on the effect of supplemental online instruction on student achievement to contribute to a larger body of research on the topic.

Although online learning is increasingly popular in U.S schools, few rigorous studies have been conducted on the effect of online learning programs on student outcomes in K-12 education. Four of the five K-12 studies included in the meta-analysis by Means et al. (2009) found positive effects of online blended programs on student achievement on researcher-developed assessments in algebra, history, and science (Long & Jennings; O'Dwyer, Carey, & Kleiman, 2007; Sun, Lin & Yu, 2008). On the other hand, recent experimental studies of other computer-based programs that were not delivered online (i.e., not web based) failed to detect positive effects of these programs on achievement on standardized tests in reading, language, and mathematics (Rouse & Krueger, 2004; Dynarski et al., 2007; Borman, Benson, & Overman, 2008; Campuzanu et al., 2009).

Two studies evaluating the effect of Reasoning Mind reached different conclusions. Using RCTs comparing Reasoning Mind with traditional mathematics programs, Weber (2003) found statistically significant effects on the Texas Assessment of Knowledge and Skills (TAKS) mathematics scores for seventh-graders in one school, while Waxman and Houston (2008) found significant effects on fifth-graders in three schools on mathematics test developed by Reasoning Mind, but not on mathematics achievement on the TAKS. There have been no prior experimental or quasi-experimental studies on the effects of DreamBox Learning.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

This study provides evidence of the impact of the two online programs as implemented by an elementary charter school network. The primary research questions are as follows:

1. What impact does supplemental online mathematics instruction (DreamBox and Reasoning Mind) have on students' mathematics learning by the end of one semester?
2. Do effects differ for students with different characteristics (i.e., English learner status, grade level, pretest scores, participation in Response to Intervention [RTI])?

Through these research questions we intend to estimate the general impact of the supplemental online mathematics instruction as well as identify sub-populations for whom the programs are most effective.

Setting:
Description of the research location.

The randomized controlled trial took place in three schools within an elementary charter school network in the San Francisco Bay Area that serves mainly low-income and minority students. These three schools had a total enrollment of more than 1,300 students.

Population / Participants / Subjects:
Description of the participants in the study: who, how many, key features, or characteristics.

This study includes a total of 1,255 students from kindergarten through fifth grade in the three schools. Among them nearly 90% are minority students, about 85% are English learners, and nearly 90% students are eligible for the federal free or reduced price meals (FRPM) program. The DreamBox experiment included 583 students in kindergarten and first grade, while the Reasoning Mind experiment included 672 students in grades 2 through 5.

Intervention / Program / Practice:
Description of the intervention, program, or practice, including details of administration and duration.
For Track 2, this may include the development and validation of a measurement instrument.

Both DreamBox and Reasoning Mind provide adaptive learning environments that tailor instruction to student needs and provide feedback to facilitating student learning. The DreamBox mathematics curriculum is based on the National Council of Teachers of Mathematics (NCTM) standards and aligned with Common Core State Standards. It focuses on learning numbers and operations, place value, and number sense. The Reasoning Mind mathematics curriculum is based on the Russian curriculum, which emphasizes in-depth understanding of arithmetic and the early introduction of algebraic notions.[2] Overall, the topics covered by both DreamBox and Reasoning Mind align with the NWEA MAP assessment; however, the extent of alignment may vary by individual student because of the adaptive nature of the curriculum.

DreamBox generates information on program use (e.g., notifications of students who are struggling with a concept or unit or working inefficiently in the program) and student progress (proficiency and growth), but does not prescribe a specific role for teachers. DreamBox Learning recommends students spend a minimum of 90 minutes per week on the program. On the other hand, Reasoning Mind is designed to be implemented in a classroom setting with a credentialed teacher. To support teachers to implement this model, Reasoning Mind provides training and ongoing support for teachers. Moreover, Reasoning Mind's design calls for 180 minutes a week of instruction (implemented as a supplemental program) with second and third graders and 450 minutes per week (implemented in an integrated fashion) with fourth and fifth graders.[3]

At the schools in the charter school network, students move between traditional classrooms and the school's "learning lab"—a computer lab that accommodates two classes of students at a time. The labs are run by lab coordinators (not teachers), who are non-credentialed, hourly staff. Some

[2] The Russian curriculum also serves as the basis for curricula in several countries with good mathematics performance, including China and Singapore.
[3] Personal communication with Reasoning Mind staff on June 20, 2011.

lab coordinators received the Reasoning Mind training while others were trained by an in-house online learning specialist. Students typically were scheduled to spend 40 minutes per day (up to 200 minutes per week) in the Learning Lab (though the range of scheduled time is from 20 to 40 minutes per day). These practices met the basic requirements of DreamBox, but not Reasoning Mind.

Research Design:
Description of the research design.

We conducted a randomized controlled trial in which students were randomly assigned to one of two conditions: (1) online mathematics instruction supplementing face-to-face mathematics instruction (treatment), or (2) face-to-face mathematics instruction only (control). We randomly assigned individual students separately within each school and by each grade level at a 4 to 1 ratio to the treatment and control groups.

All students received approximately 100 to 110 minutes per day of face-to-face mathematics instruction. Students in the treatment group received an additional 20 to 40 minutes per day of online mathematics instruction over the course of 4 months (mid-October through mid-February), while the control students in the same class received online literacy instruction in the same computer lab. Therefore this evaluation essentially estimates the effect of supplemental online mathematics instruction versus the online literacy instruction on students' mathematics outcomes.

The charter school network administered the Northwest Evaluation Association's (NWEA) mathematics tests in September 2010 (pretest) and January/February 2011 (posttest) to students included in the experiment. Our analysis included both the overall NWEA mathematics scores and subtest scores for problem solving, number sense, computation, measurement and geometry, statistics and probability, Algebra and function, and math reasoning. All the scores are in the RIT scale,[4] which is scaled using the Item Response Theory (IRT) and has the same meaning regardless of the grade of the student. Table 1 and 2 list the pre- and posttest scores respectively for treatment and control students in each program.

(Please insert Table 1 and Table 2 here)

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.
For Track 2, this may include the use of existing datasets.

We collected student demographic information and September and January/February scores on the NWEA mathematics test. We also collected usage data from the two online programs that provide the hours students spent on the programs over the course of the experiment. In addition, we requested computer lab schedules in these schools and school calendars, from which we calculated scheduled participation time.

Usage data reveal a considerable amount of treatment crossover and variation—that is, control students participating in the program and treatment students receiving varied dosage. In all three

[4] The RIT Scale is a curriculum scale that uses individual item difficulty values to estimate student achievement. For more information, see http://www.nwea.org/support/article/532/rit-scale

schools, some low-achieving students, regardless of their treatment assignment, were assigned to a Response to Intervention (RtI) program, where they received literacy tutoring as well as about 45 minutes of DreamBox or Reasoning Mind each day. Therefore treatment students in the RtI program on average spent more time on the online learning programs than those who were not included in the program and control students in the RtI program had regular access to the online curricula. In addition to the RtI students in the control group, other control students also had access to the online learning programs in the computer lab and, though it was not intended to be part of their instructional day, some of them did log into the programs. Tables 3 and 4 show the hours of program participation for treatment and control students by RtI status for DreamBox and Reasoning Mind separately. Even taking RtI into account, the variation in program participation among treatment students is quite large.

(Please insert Table 3 and Table 4 here)

We conducted two types of analysis to examine the effects of the online mathematics programs. One is an intent-to-treat (ITT) analysis where we studied the effect of being assigned to the treatment group regardless of each student's actual time spent on the program. We estimated the ITT effect on the posttest achievement adjusting for students' demographic background, pretest scores, RtI status, grade level, and school fixed effects.[5] We also looked at the interaction between treatment and pretest score, gender, FRPM, RtI status, grade level, and school fixed effects to examine whether the online programs have differential effects on student subgroups.

The ITT analysis offers an unbiased estimate between the treatment and control groups but it tends to underestimate the effect of getting the treatment since some control students actually got the treatment while some treatment students did not. Therefore we also conducted a treatment-on-treated (TOT) analysis using an instrumental variable (IV) approach, where we applied a two-stage least squares regression, using treatment assignment as the instrument to predict the actual hours a student participated in the program and then estimating the effect of the predicted program hours on the outcomes. The effect of predicted participation hours, unlike actual hours students spent on the programs, is not subject to selection bias; thus, we could obtain an unbiased estimate of the effect of participation. Not detailed here, the methods we used met the conditions proposed by Angrist et al. (1996) for the application of the IV approach.

Findings / Results:

Description of the main findings with specific details.

Students in the DreamBox treatment group scored an average of 2.3 points higher on the NWEA overall mathematics test than similar students in the control group; this difference translates into an effect size of .14. This difference also translates to an improvement index of 5.5 percentile points, which suggests that being assigned to the treatment group would have led to a 5.5 point increase in the percentile rank for the average (50th percentile) student in the control group. Students in the treatment group scored an average of 2.9 points higher on the measurement and geometry subtest than their peers in the control group; this difference translates into an effect size of .16. This difference also translates to an improvement index of 6.4 percentile points, suggesting that being assigned to the treatment group would have led to a 6.4 point increase in the percentile rank for the average (50th percentile) student in the control group. Although we

[5] We also posited hierarchical models with classroom and student levels, with treatment status at the student level. The results are very similar to that from the OLS regression and are not presented in this paper.

found no statistically significant effects on the problem solving, number sense, computation, or statistics and probability subtests, the effects all have a positive sign, suggesting that DreamBox improved student math achievement in a comprehensive way.

We did not find a statistically significant effect of Reasoning Mind on either the NWEA overall or any of the subtest scores. In addition, we did not find statistically significant and consistent differential DreamBox or Reasoning Mind effects for student subgroups. Tables 5 to 8 present the results of the regression models for DreamBox and Reasoning Mind, respectively.

(Please insert Table 5 to Table 8 here)

Consistent with the ITT findings, we found significant TOT effects of DreamBox usage hours on NWEA overall mathematics test score as well as on the measurement and geometry subtest score, but not on the problem solving, number sense, and statistics and probability subtests. Likewise, we did not find a statistically significant effect of hours spent on Reasoning Mind on any of the NWEA scores. Tables 9 and 10 present the results of the two-stage least squares regression models for DreamBox and Reasoning Mind, respectively.

(Please insert Table 9 and Table 10 here)

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

This study's positive findings about the effects of DreamBox instruction are likely to fuel the sense of optimism about the promise of online learning, especially in light of the relatively modest treatment. In interpreting these findings, we urge educators and policymakers to keep in mind a basic principle of scientific research—that research findings contribute to the ongoing refinement of hypotheses but do not represent a conclusion. Positive results merit continued and even expanded use, but ongoing evaluation is needed to build a body of evidence, especially as interventions are implemented in varied ways in diverse settings. We also urge readers to consider that there may be a variety of reasons that we did not find an effect of Reasoning Mind on NWEA mathematics scores. It may reflect true shortcomings in the effectiveness (power to engage students and enhance learning) of the program. Other possible explanations may be that implementation of Reasoning Mind was not optimal (e.g., Reasoning Mind is designed to be used in the classroom with teachers guiding instruction) or that Reasoning Mind may require more instructional time (i.e., minutes per week) or take longer than 4 months to take effect.

In fact, this study only examined the effects of using the DreamBox and Reasoning Mind programs for a short period of time. Using the program for a longer period of time may have led to different effects. Moreover, because we only examined the short-term effects of the program, we do not know whether or for how long the estimated positive DreamBox effect would persist and whether Reasoning Mind might have an effect on longer-term outcomes. Further follow-up of the experiment would be needed to address these questions.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Borman, G.D., Benson, J, & Overman, L. (2009). A randomized field trial of the Fast ForWord Language computer-based training program. *Educational Evaluation and Policy Analysis 31*: 82-106.

Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). *Effectiveness of reading and mathematics software products: Findings from two student cohorts* (NCEE 2009–4041). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., et al. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort.* Washington, D.C.: U.S. Department of Education.

Horn, M. B. & Staker, H.. (2011) *The rise of K-12 blended learning.* Mountain View, CA: Innosight Institute.

Long, M., & Jennings, H. (2005). *"Does it work?": The impact of technology and professional development on student achievement.* Calverton, Md.: Macro International.

Means, B., Toyama,Y., Murphy. R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online-learning studies.* Washington, D.C.: U.S. Department of Education.

O'Dwyer, L. M., Carey, R., & Kleiman, K. (2007). A study of the effectiveness of the Louisiana Algebra I online course. *Journal of Research on Technology in Education 39* (3):289–306.

Rocketship Education (2011). Mission statement. Retrieved from http://rsed.org/index.php?page=mission

Rockman et al. (2007). *ED PACE final report*. Submitted to the West Virginia Department of Education. San Francisco: Author. Retrieved from www.rockman.com/projects/146.ies.edpace/finalreport.

Rouse, C. E. & Krueger, A. B. (2004). Putting computerized instruction to the test: a randomized evaluation of a "scientifically based" reading program. *Economics of Education Review 23*(4): 323-338.

Sun, K., Lin, Y., & Yu, C (2008). A study on learning effect among different learning styles in a Web-based lab of science for elementary school students. *Computers & Education 50* (4):1411–22.

U.S. Department of Education, Office of Educational Technology (2010). *Transforming American education: learning powered by technology. National education technology plan 2010.* Washington, DC: Author.

Weber, W.A. (2003). An evaluation of the Reasoning Mind Pilot Program at Hogg Middle School. Retrieved from http://www.reasoningmind.org/pdf/Dr_Weber_Report_2003.pdf

Waxman, H. C., & Houston, W. R. (2008). An evaluation of the 2006-07 Reasoning Mind Program. Retrieved from http://www.reasoningmind.org/pdf/Waxman-Houston_Report_2007.pdf

Appendix B. Tables and Figures
Not included in page count.

Table 1
Pre and Post NWEA Mathematics Test Scores for Students in the DreamBox Experiment

	Treatment					Control				
	Pretest			Post-test		Pretest			Post-test	
	N	Mean	SD	Mean	SD	N	Mean	SD	Mean	SD
Math Overall	446	146.0	18.0	159.0	16.6	111	144.7	15.0	156.2	15.1
Problem Solving	444	147.0	19.3	161.4	16.3	109	144.7	17.1	159.8	15.2
Number Sense	444	146.9	20.0	159.6	18.9	109	143.4	16.6	157.0	17.2
Computation	438	147.5	22.4	163.0	20.7	108	147.0	19.8	158.8	19.5
Measurement and Geometry	441	144.5	18.9	155.5	18.3	109	144.8	18.4	151.8	18.1
Statistics and Probability	443	145.5	19.3	156.3	18.9	109	145.1	15.6	154.1	17.6

Table 2
Pre and Post NWEA Mathematics Test Scores for Students in the Reasoning Mind Experiment

	Treatment					Control				
	Pretest			Post-test		Pretest			Post-test	
	N	Mean	SD	Mean	SD	N	Mean	SD	Mean	SD
Math Overall	521	185.8	17.5	194.2	15.4	130	187.3	18.3	195.8	15.9
Number Sense	518	187.1	20.2	198.2	16.5	130	188.9	21.8	200.7	18.9
Algebra and Function	518	186.7	18.1	196.0	16.7	130	187.2	19.6	196.2	17.3
Measurement and Geometry	518	187.0	19.1	194.0	16.2	130	188.3	18.6	196.5	16.9
Statistics and Probability	518	185.3	18.0	191.8	16.4	130	187.3	18.0	192.9	16.0
Math Reasoning	518	183.3	18.8	191.4	17.7	130	184.5	19.4	192.8	17.5

Table 3
DreamBox Participation Hours by Treatment Condition and RtI Status

		N	Mean	SD
Treatment				
	Overall	446	21.8	6.7
	Non-RtI	404	21.3	6.3
	RtI	42	26.5	8.2
Control				
	Overall	111	5.1	3.4
	Non-RtI	100	4.8	3.2
	RtI	11	8.0	4.1

Table 4
Reasoning Mind Participation Hours by Treatment Condition and RtI Status

		N	Mean	SD
Treatment				
	Overall	521	20.7	9.8
	Non-RtI	365	18.7	7.6
	RtI	156	25.4	12.5
Control				
	Overall	130	3.0	7.3
	Non-RtI	91	0.7	2.5
	RtI	39	8.3	11.2

Table 5
Regression Results for the ITT Effects of DreamBox on NWEA Math Scores

	Math Overall (*N* = 552)		Problem Solving (*N* = 549)		Number Sense (*N* = 549)		Computation (*N* = 543)		Measurement and Geometry (*N* = 546)		Statistics and Probability (*N* = 550)	
Intercept	59.07		81.98		64.78		72.81		48.60		49.94	
	(4.26)		(5.63)		(6.25)		(7.25)		(6.24)		(6.92)	
Pretest math overall	0.68	***	0.54	***	0.73	***	0.61	***	0.71	***	0.60	***
	(0.03)		(0.06)		(0.07)		(0.07)		(0.07)		(0.07)	
Pretest to outcome			0.01		-0.06		-0.04		0.04		0.11	
			(0.04)		(0.05)		(0.06)		(0.06)		(0.06)	
Treatment	**2.30**	**	**1.02**		**1.53**		**2.68**		**2.91**	*	**2.20**	
	(0.83)		**(1.11)**		**(1.23)**		**(1.41)**		**(1.23)**		**(1.36)**	
Female	0.34		1.29		-0.44		0.64		0.54		-0.09	
	(0.67)		(0.89)		(0.98)		(1.14)		(0.99)		(1.09)	
Hispanic	-2.58	*	-3.83	**	-4.22	**	-5.59	**	-2.52		1.79	
	(1.06)		(1.41)		(1.55)		(1.80)		(1.57)		(1.73)	
English learner	0.59		-2.22		1.56		2.43		0.13		0.10	
	(0.89)		(1.18)		(1.30)		(1.50)		(1.31)		(1.43)	
FRPM	-2.52	*	0.15		-2.91		-3.52		-3.38	*	-1.91	
	(1.14)		(1.51)		(1.65)		(1.92)		(1.67)		(1.83)	
Special education	-2.24		-0.82		-1.52		-2.73		-2.24		-3.05	
	(1.64)		(2.17)		(2.44)		(2.82)		(2.40)		(2.66)	
RtI	-4.51	***	-2.82		-6.33	****	-3.00		-5.20	**	-5.42	**
	(1.23)		(1.62)		(1.79)		(2.08)		(1.79)		(1.99)	
Grade 1	6.12	***	6.50	***	7.71	***	14.16	***	1.35		3.14	
	(1.03)		(1.36)		(1.52)		(1.84)		(1.53)		(1.68)	
School A	-0.16		-0.70		-2.21		3.72	*	0.00		-1.17	
	(0.87)		(1.15)		(1.27)		(1.48)		(1.28)		(1.41)	
School B	-0.73		-0.78		-3.14	**	4.74	***	-2.33	*	-0.78	
	(0.80)		(1.06)		(1.17)		(1.36)		(1.18)		(1.29)	
R^2	0.78		0.60		0.63		0.60		0.61		0.55	

*p < .05, **p < .01, ***p < .001

Table 6
Regression Results for the ITT Effects of DreamBox on NWEA Math Scores with Subgroup Interactions

	Math Overall (*N* = 552)		Problem Solving (*N* = 551)		Number Sense (*N* = 550)		Computation (*N* = 543)		Measurement and Geometry (*N* = 549)		Statistics and Probability (*N* = 549)	
Intercept	57.40	***	78.65	***	80.34	***	57.70	**	67.75	***	57.24	***
	(10.77)		(14.32)		(15.68)		(18.08)		(15.92)		(17.10)	
Pretest math overall	0.70	***	0.56	***	0.62	***	0.71	***	0.56	***	0.58	***
	(0.07)		(0.10)		(0.13)		(0.14)		(0.13)		(0.13)	
Pretest to outcome			0.00		-0.06		-0.04		0.07		0.10	
			(0.05)		(0.05)		(0.06)		(0.06)		(0.06)	
Treatment	3.82		5.13		-15.50		20.09		-19.97		-7.67	
	(11.60)		(15.47)		(16.91)		(19.53)		(17.22)		(18.47)	
Female	-0.28		-1.21		-1.74		-1.21		0.48		-1.78	
	(1.52)		(2.03)		(2.23)		(2.54)		(2.28)		(2.45)	
Hispanic	-2.60	*	-3.74	**	-3.94	**	-5.44	**	-2.99		1.97	
	(1.07)		(1.45)		(1.58)		(1.80)		(1.62)		(1.74)	
English learner	-1.88		-3.41		-2.87		-2.95		1.25		-1.03	
	(1.95)		(2.59)		(2.83)		(3.28)		(2.93)		(3.14)	
FRPM	-2.42		1.87		-0.71		2.22		-7.51		-4.64	
	(2.61)		(3.49)		(3.81)		(4.35)		(3.91)		(4.21)	
Special education	-2.30		-1.03		-1.65		-2.90		-2.25		-3.29	
	(1.64)		(2.21)		(2.46)		(2.80)		(2.47)		(2.66)	
RtI	-0.94		-0.07		-2.35		1.69		-3.33		-7.64	
	(2.81)		(3.77)		(4.15)		(4.72)		(4.22)		(4.52)	
Grade 1	6.50	**	8.33	**	10.90	**	14.80	***	4.81		4.00	
	(2.37)		(3.14)		(3.50)		(3.93)		(3.51)		(3.75)	
School A	2.01		0.98		0.88		4.60		1.36		1.86	
	(1.96)		(2.63)		(2.87)		(3.30)		(2.96)		(3.17)	
School B	0.23		-0.03		-2.48		5.73		-5.46	*	1.58	
	(1.82)		(2.46)		(2.72)		(3.06)		(2.77)		(2.95)	
Treatment*Pretest math overall	-0.02		-0.01		0.12		-0.11		0.13		0.05	
	(0.08)		(0.11)		(0.12)		(0.14)		(0.12)		(0.13)	
Treatment*Female	0.77		3.43		1.39		2.15		-0.45		2.31	
	(1.69)		(2.27)		(2.49)		(2.83)		(2.54)		(2.74)	
Treatment*Eng lrnr	3.19		0.87		5.16		6.89		(0.56)		1.46	
	(2.19)		(2.92)		(3.18)		(3.67)		(3.29)		(3.53)	
Treatment*FRPM	-0.09		-1.99		-2.57		-7.03		5.75		3.51	
	(2.89)		(3.89)		(4.23)		(4.83)		(4.34)		(4.66)	
Treatment*RtI	-4.51		-3.24		-5.28		-5.96		-2.76		2.85	
	(3.12)		(4.19)		(4.61)		(5.25)		(4.69)		(5.02)	
Treatment*Grade 1	-0.57		-2.48		-3.83		-1.20		-3.25		-1.43	
	(2.61)		(3.50)		(3.85)		(4.31)		(3.86)		(4.16)	
Treatment*Sch A	-2.71		-2.45		-3.67		-1.09		-1.28		-4.09	
	(2.19)		(2.94)		(3.21)		(3.69)		(3.29)		(3.54)	
Treatment*Sch B	-1.09		-1.36		-0.76		-0.64		4.01		-3.17	
	(2.03)		(2.74)		(3.03)		(3.42)		(3.07)		(3.30)	
R^2	0.781		0.594		0.633		0.612		0.606		0.558	

*p < .05, **p < .01, ***p < .001

Table 7
Regression Results for the ITT Effects of Reasoning Mind on NWEA Mathematics Scores

	Math Overall (*N* = 646)		Number Sense (*N* = 644)		Algebra and Function (*N* = 645)		Measurement and Geometry (*N* = 644)		Statistics and Probability (*N* = 640)		Math Reasoning (*N* = 643)	
Intercept	75.94		81.23		79.21		84.58		82.17		64.03	
	(5.23)		(6.84)		(7.06)		(6.85)		(7.05)		(7.02)	
Pretest math overall	0.66	***	0.74	***	0.56	***	0.57	***	0.53	***	0.66	***
	(0.03)		(0.07)		(0.07)		(0.06)		(0.07)		(0.06)	
Pretest of outcome			-0.11		0.09		0.04		0.08		0.05	
			(0.05)		(0.06)		(0.06)		(0.06)		(0.06)	
Treatment	**-0.47**		**-1.74**		**0.60**		**-1.82**		**-0.12**		**-0.02**	
	(0.73)		**(0.94)**		**(0.99)**		**(0.95)**		**(0.98)**		**(0.97)**	
Female	0.32		0.19		0.11		0.45		0.08		0.53	
	(0.58)		(0.76)		(0.79)		(0.77)		(0.78)		(0.78)	
Hispanic	-1.63		-0.72		-2.01		-0.52		-2.15		-3.96	**
	(0.96)		(1.26)		(1.32)		(1.27)		(1.30)		(1.29)	
English learner	-0.51		-0.24		-0.37		-0.49		-0.13		0.03	
	(0.69)		(0.89)		(0.93)		(0.90)		(0.92)		(0.92)	
FRPM	-0.97		0.59		-0.94		-2.01		-1.42		-1.24	
	(1.20)		(1.56)		(1.64)		(1.58)		(1.62)		(1.61)	
Special education	-5.88	***	-5.26	**	-5.17	**	-6.52	***	-5.97	***	-6.15	***
	(1.32)		(1.72)		(1.79)		(1.74)		(1.81)		(1.77)	
RtI	-2.69	***	-3.24	**	-2.63	*	-2.90	**	-2.89	**	-3.44	***
	(0.76)		(0.99)		(1.03)		(1.00)		(1.02)		(1.02)	
Grade 3	2.42	**	2.23	*	1.93		1.58		3.90	***	2.76	**
	(0.77)		(1.00)		(1.04)		(1.01)		(1.04)		(1.03)	
Grade 4	4.11	***	6.25	***	5.24	***	3.90	**	2.65		3.18	*
	(1.02)		(1.33)		(1.39)		(1.35)		(1.39)		(1.37)	
Grade 5	5.55	***	7.39	***	6.30	***	7.69	***	6.38	***	1.58	
	(1.40)		(1.82)		(1.91)		(1.85)		(1.92)		(1.88)	
School A	-1.46		-0.36		-3.60	**	-1.79		-1.98		-0.63	
	(0.88)		(1.15)		(1.20)		(1.16)		(1.19)		(1.19)	
School B	-2.25	**	-1.70		-3.17	**	-1.89		-2.78	*	-2.59	*
	(0.81)		(1.05)		(1.10)		(1.07)		(1.09)		(1.09)	
R^2	0.78		0.68		0.65		0.66		0.63		0.69	

*p < .05, **p < .01, ***p < .001

Table 8
Regression Results for the ITT Effects of Reasoning Mind on NWEA Mathematics Scores with Subgroup Interactions

	Math Overall (*N* = 646)		Number Sense (*N* = 644)		Algebra and Function (*N* = 644)		Measurement and Geometry (*N* = 643)		Statistics and Probability (*N* = 642)		Math Reasoning (*N* = 643)	
Intercept	80.69	***	73.90	***	97.99	***	105.20	***	88.95	***	66.19	***
	(11.22)		(14.28)		(14.94)		(14.37)		(15.23)		(14.94)	
Pretest math general	0.62	***	0.77	***	0.47	***	0.45	***	0.51	***	0.61	***
	(0.06)		(0.10)		(0.10)		(0.09)		(0.10)		(0.10)	
Pretest of outcome			-0.11	*	0.07		0.04		0.09		0.07	
			(0.05)		(0.06)		(0.06)		(0.06)		(0.06)	
Treatment	-4.82		10.16		-20.20		-26.66		-10.10		-2.45	
	(12.52)		(15.96)		(16.72)		(16.08)		(16.99)		(16.69)	
Female	0.89		-0.89		1.98		0.44		-0.12		2.31	
	(1.33)		(1.72)		(1.80)		(1.73)		(1.80)		(1.77)	
Hispanic	-1.54		-0.92		-2.09		-0.46		-1.36		-3.87	**
	(0.97)		(1.26)		(1.32)		(1.26)		(1.31)		(1.29)	
English learner	2.69		1.61		1.09		2.11		3.17		3.60	
	(1.54)		(1.97)		(2.05)		(1.98)		(2.08)		(2.05)	
FRPM	-2.65		-0.80		-1.07		-4.90		-6.99	*	1.76	
	(2.64)		(3.41)		(3.56)		(3.43)		(3.57)		(3.51)	
Special education	-5.98	***	-5.55	***	-5.38	**	-6.71	***	-5.76	**	-6.21	***
	(1.33)		(1.72)		(1.80)		(1.74)		(1.84)		(1.77)	
RtI	-4.47	**	-1.62		-5.91	*	-6.51	**	-5.54	*	-7.15	**
	(1.76)		(2.24)		(2.34)		(2.26)		(2.38)		(2.34)	
Grade 3	2.41		2.43		1.83		2.80		2.64		3.27	
	(1.82)		(2.36)		(2.46)		(2.37)		(2.47)		(2.43)	
Grade 4	2.55		3.54		3.55		4.70		-0.17		1.84	
	(2.26)		(2.93)		(3.06)		(2.95)		(3.06)		(3.01)	
Grade 5	10.28	**	12.00	**	13.89	**	19.22	***	7.46		5.80	
	(3.43)		(4.44)		(4.63)		(4.59)		(4.65)		(4.57)	
School A	-1.71		5.18		-5.30		-0.12		-2.66		-3.27	
	(2.05)		(2.66)		(2.77)		(2.67)		(2.78)		(2.73)	
School B	-0.82		0.76		-1.46		3.21		-2.77		-1.53	
	(1.89)		(2.44)		(2.55)		(2.45)		(2.57)		(2.51)	
Treatment*Pretest math overall	0.03		-0.05		0.12		0.15		0.02		0.04	
	(0.07)		(0.08)		(0.09)		(0.08)		(0.09)		(0.09)	
Treatment*Female	-0.60		1.31		-2.04		0.13		0.21		-2.22	
	(1.48)		(1.91)		(2.00)		(1.93)		(2.00)		(1.97)	
Treatment*ELL	(4.02)	*	(2.42)		(1.76)		(3.16)		(3.98)		(4.30)	
	(1.72)		(2.21)		(2.30)		(2.22)		(2.34)		(2.29)	
Treatment*FRPM	2.02		1.57		0.07		3.67		6.85		-3.75	
	(2.95)		(3.82)		(3.98)		(3.84)		(3.99)		(3.92)	
Treatment*RtI	2.10		-2.00		3.89		4.44		3.34		4.38	
	(1.95)		(2.50)		(2.60)		(2.52)		(2.64)		(2.61)	
Treatment*Grade 3	0.27		0.09		0.38		-1.23		1.71		-0.75	

	(2.00)	(2.60)	(2.71)	(2.61)	(2.72)	(2.68)
Treatment*Grade 4	2.06	3.74	2.13	-1.00	3.36	1.36
	(2.53)	(3.28)	(3.43)	(3.30)	(3.45)	(3.37)
Treatment*Grade 5	-5.40	-5.15	-8.74	-12.97 **	-2.00	-5.10
	(3.73)	(4.82)	(5.03)	(4.97)	(5.06)	(4.97)
Treatment*School A	0.18	-6.80 *	2.10	-2.09	0.59	3.05
	(2.26)	(2.93)	(3.06)	(2.95)	(3.08)	(3.01)
Treatment*School B	-1.79	-2.90	-2.19	-6.36 *	-0.11	-1.27
	(2.09)	(2.70)	(2.82)	(2.71)	(2.85)	(2.78)
R^2	0.783	0.69	0.659	0.672	0.629	0.702

*p < .05, **p < .01, ***p < .001

Table 9
Results for Two-Stage Least Squares Regression of the TOT Effects of DreamBox Participation on NWEA Math Scores

	Math Overall (*N* = 552)		Problem Solving (*N* =549)		Number Sense (*N* =549)		Computation (*N* = 543)		Measurement and Geometry (*N* =546)		Statistics and Probability (*N* = 550)	
Intercept	57.96		81.49		64.11		71.59		47.22		49.06	
	(4.27)		(5.69)		(6.34)		(7.29)		(6.30)		(6.99)	
DreamBox participation hours	**0.14**	**	**0.06**		**0.09**		**0.16**		**0.17**	*	**0.13**	
	(0.05)		**(0.07)**		**(0.07)**		**(0.08)**		**(0.07)**		**(0.08)**	
Pretest math overall	0.69	***	0.55	***	0.73	***	0.61		0.72	***	0.62	***
	(0.05)		(0.06)		(0.07)		(0.07)		(0.07)		(0.07)	
Pretest of outcome			0.01		-0.06		-0.04		0.03		0.10	
			(0.04)		(0.05)		(0.06)		(0.06)		(0.06)	
Female	0.33		1.29		-0.43		0.65		0.55		-0.09	
	(0.66)		(0.89)		(0.98)		(1.13)		(0.98)		(1.08)	
Hispanic	-2.62	*	-3.85	**	-4.25	**	-5.64	**	-2.59		1.73	
	(1.05)		(1.40)		(1.55)		(1.77)		(1.56)		(1.72)	
English learner	0.62		-2.20		1.58		2.46		0.18		0.12	
	(0.87)		(1.17)		(1.29)		(1.49)		(1.30)		(1.43)	
FRPM	-2.47	*	0.16		-2.87		-3.44		-3.26	*	-1.83	
	(1.13)		(1.51)		(1.65)		(1.91)		(1.66)		(1.83)	
Special education	-1.91		-0.67		-1.29		-2.43		-1.82		-2.76	
	(1.61)		(2.15)		(2.43)		(2.78)		(2.38)		(2.65)	
RtI	-4.95	***	-3.02		-6.62	***	-3.55		-5.75	***	-5.85	**
	(1.22)		(1.62)		(1.80)		(2.08)		(1.80)		(1.99)	
Grade 1	5.62	***	6.29	***	7.40	***	13.56	***	0.74		2.68	
	(1.03)		(1.38)		(1.54)		(1.85)		(1.54)		(1.69)	
School A	-0.23		-0.74		-2.26		3.66	*	-0.08		-1.26	
	(0.85)		(1.15)		(1.27)		(1.47)		(1.27)		(1.41)	
School B	-1.12		-0.96		-3.40	**	4.27	**	-2.84	**	-1.15	
	(0.79)		(1.07)		(1.18)		(1.36)		(1.19)		(1.31)	
R^2	0.78		0.59		0.63		0.60		0.61		0.54	

*p < .05, **p < .01, ***p < .001

Table 10
Results for Two-Stage Least Squares Regression of the TOT Effects of Reasoning Mind Participation on NWEA Mathematics Scores

	Math Overall (*N* =646)		Number Sense (*N* =644)		Algebra and Function (*N* = 645)		Measurement and Geometry (*N* = 644)		Statistics and Probability (*N* = 640)		Math Reasoning (N=643)	
Intercept	75.33		79.17		79.97		82.27		82.03		64.01	
	(5.20)		(6.87)		(6.97)		(6.85)		(6.98)		(6.93)	
Reasoning Mind participation hours	**-0.03**		**-0.10**		**0.03**		**-0.10**		**-0.01**		**0.00**	
	(0.04)		**(0.05)**		**(0.05)**		**(0.05)**		**(0.05)**		**(0.05)**	
Pretest math overall	0.66	***	0.75	***	0.56	***	0.59	***	0.53	***	0.66	***
	(0.03)		(0.07)		(0.07)		(0.06)		(0.07)		(0.07)	
Pretest of outcome			-0.10		0.09		0.04		0.08		0.05	
			(0.06)		(0.06)		(0.06)		(0.06)		(0.06)	
Female	0.312		0.16		0.12		0.43		0.08		0.53	
	-0.585		(0.77)		(0.79)		(0.78)		(0.78)		(0.78)	
Hispanic	-1.60		-0.63		-2.05		-0.42		-2.15		-3.96	**
	(0.97)		(1.28)		(1.31)		(1.28)		(1.30)		(1.29)	
English learner	-0.50		-0.23		-0.37		-0.46		-0.13		0.03	
	(0.69)		(0.90)		(0.92)		(0.91)		(0.92)		(0.92)	
FRPM	-0.99		0.52		-0.91		-2.09		-1.42		-1.24	
	(1.21)		(1.59)		(1.63)		(1.60)		(1.62)		(1.61)	
Special education	-6.04	***	-5.85	***	-4.96	**	-7.12	***	-6.01	***	-6.15	***
	(1.34)		(1.75)		(1.80)		(1.77)		(1.82)		(1.79)	
RtI	-2.46	**	-2.37	*	-2.92	**	-1.98		-2.83	*	-3.43	**
	(0.84)		(1.09)		(1.12)		(1.11)		(1.12)		(1.12)	
Grade 3	2.39	**	2.10	*	1.97		1.45		3.89	***	2.76	**
	(0.77)		(1.02)		(1.04)		(1.03)		(1.04)		(1.03)	
Grade 4	4.05	***	6.04	***	5.30	***	3.68	**	2.64		3.18	*
	(1.03)		(1.35)		(1.39)		(1.36)		(1.40)		(1.37)	
Grade 5	5.40	***	6.82	***	6.50	***	7.08	***	6.35	***	1.57	
	(1.43)		(1.87)		(1.92)		(1.88)		(1.94)		(1.90)	
School A	-1.49		-0.48		-3.56	**	-1.89		-1.99		-0.63	
	(0.88)		(1.17)		(1.20)		(1.18)		(1.19)		(1.19)	
School B	-2.26	**	-1.75		-3.14	**	-1.94		-2.78	*	-2.59	*
	(0.81)		(1.07)		(1.10)		(1.08)		(1.09)		(1.09)	
R^2	0.77		0.67		0.65		0.64		0.62		0.68	

*p < .05, **p < .01, ***p < .001