5-1-2017

The Impact of OGAP on Elementary Math Teacher Knowledge and Student Achievement

Jonathan A. Supovitz
University of Pennsylvania

Janine Remillard
University of Pennsylvania

Robert A. Nathenson
University of Pennsylvania

The Impact of OGAP on Elementary Math Teacher Knowledge and Student Achievement

Abstract

The Ongoing Assessment Project (OGAP) is a learning trajectory-oriented formative assessment program that develops teachers' abilities to understand and apply research-based developmental trajectories in math content areas to deepen their thinking about their students. In OGAP, teachers learn to use a learning progression framework to continually assess and adapt their instruction to students' developing understanding, aiming to move them towards more sophisticated strategies in a range of multiplicative contexts. For this reason, OGAP puts a premium on students' precision of answer (including correctness and unit labeling) and sophistication of solution response.

In this study we examine the multi-year impacts of OGAP on grades 3-5 student correctness and solution sophistication in multiplication on an open-ended assessment created by the Consortium for Policy Research in Education (CPRE) as part of a randomized experimental study of OGAP in Philadelphia schools. In order to assess the intervention's impact on student learning in both correctness and sophistication, the research team developed an assessment measure with three vertically-equated grade-specific forms composed of open-ended items. The assessment asked students to show their work to allow for analysis of their correctness, strategies, and errors.

The results show strong and consistent first year effects on student correctness and solution sophistication multiplication outcomes in all three grades that were assessed. However, these results did not persist during the second year of OGAP treatment, which focused on fraction, when controlling for end of first year results. When examining the second-year multiplication results using the baseline measure, the treatment impacts were present, reinforcing the strength of the first year effects. The next step is to examine year 2 effects in fractions, which was the focus of the second year of OGAP professional development. Additionally, since student and teacher turnover are manifest in Philadelphia, and consequently both students and teachers had different levels of exposure to OGAP, additional analyses are needed to incorporate student and teacher levels of exposure and implementation of OGAP into the models, to disentangle results by level of treatment.

Keywords

OGAP, NSF, Ongoing Assessment

Disciplines

Educational Assessment, Evaluation, and Research | Educational Methods | Elementary Education | Science and Mathematics Education

Two Year Impacts of the Ongoing Assessment Project on Student Multiplicative Reasoning in Philadelphia, PA[1]

Jonathan Supovitz, Caroline Ebby, Janine Remillard, Robert Nathenson
Graduate School of Education, University of Pennsylvania
Consortium for Policy Research in Education

Paper presented at the American Educational Research Association
San Antonio, Texas
May 1, 2017

DRAFT – NOT FOR DISTRIBUTION

Abstract

The Ongoing Assessment Project (OGAP) is a learning trajectory-oriented formative assessment program that develops teachers' abilities to understand and apply research-based developmental trajectories in math content areas to deepen their thinking about their students. In OGAP, teachers learn to use a learning progression framework to continually assess and adapt their instruction to students' developing understanding, aiming to move them towards more sophisticated strategies in a range of multiplicative contexts. For this reason, OGAP puts a premium on students' precision of answer (including correctness and unit labeling) and sophistication of solution response.

In this study we examine the multi-year impacts of OGAP on grades 3-5 student correctness and solution sophistication in multiplication on an open-ended assessment created by the Consortium for Policy Research in Education (CPRE) as part of a randomized experimental study of OGAP in Philadelphia schools. In order to assess the intervention's impact on student learning in both correctness and sophistication, the research team developed an assessment measure with three vertically-equated grade-specific forms composed of open-ended items. The assessment asked students to show their work to allow for analysis of their correctness, strategies, and errors.

The results show strong and consistent first year effects on student correctness and solution sophistication multiplication outcomes in all three grades that were assessed. However, these results did not persist during the second year of OGAP treatment, which focused on fraction, when controlling for end of first year results. When examining the second-year multiplication results using the baseline measure, the treatment impacts were present, reinforcing the strength of the first year effects. The next step is to examine year 2 effects in fractions, which was the focus of the second year of OGAP professional development. Additionally, since student and teacher turnover are manifest in Philadelphia, and consequently both students and teachers had different levels of exposure to OGAP, additional analyses are needed to incorporate student and

[1] This paper is based on work supported by the National Science Foundation under Grant No. DRK-12 1316527. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

teacher levels of exposure and implementation of OGAP into the models, to disentangle results by level of treatment.

Introduction

In order to effectively implement more challenging standards in classrooms, teachers must have a wealth of knowledge about subject matter, student thinking and understanding, and instructional pedagogies (Ball, Thames, & Phelps, 2008; Carpenter et al., 1989; Hill, Rowan, & Ball, 2005; Ma, 1999; Petit, Laird, & Marsden, 2010; Shulman, 1987). In particular, the focus on *learning trajectories* places new demands on teaching, as teachers must not only understand the mathematical ideas and skills embodied in the standards, but also assess where students are in the trajectory of learning those concepts and skills, and then use that information to design and enact instructional responses that supports students' movement along that trajectory.

Research suggests that teachers' skill with formative assessment is a key factor for improving student learning in mathematics (Black & Wiliam, 1998; Wiliam, 2007). In their extensive review of the literature on formative assessment, Black and Wiliam (1998) found substantial evidence linking formative assessment with higher student achievement, with typical effect sizes ranging from an impressive 0.4 to 0.7. Across these studies, formative assessment was shown to be particularly beneficial for low-performing students, which suggests that increasing teachers' skill with formative assessment has promise for closing the achievement gap.

Though clearly critical for improving student learning, using learning trajectories in formative assessment is challenging work. The topics of multiplication, division, rational numbers, ratio and proportion are all considered part of the *multiplicative conceptual field* and as such are interconnected rather than isolated topics (Vergnaud, 1994). Over the last 30 years, a wealth of research has been conducted on children's development of these concepts (e.g., Behr, Harel, Post, & Lesh, 1992; Harel & Confrey, 1994; Lamon, 2007; Steffe & Olive, 2010; Tournaire & Pulos, 1985) and more recently on the development of learning trajectories for multiplicative reasoning (e.g., Sherin & Fuson, 2005), rational number reasoning, and proportional reasoning (e.g., Confrey, 2009). These learning trajectories have the potential to provide teachers with a clear articulation of learning goals, how students' thinking should develop, and learning activities that are likely to move students along the path towards achieving those goals (Heritage, 2008; Sztajn, Confrey, Wilson, & Edginton, 2012). However, little of this research has been translated into a form that makes learning trajectories concrete and usable for teachers to use in their day-to-day instruction (Daro, Mosher & Corcoran, 2011).

Additionally, formative assessment involves more than simply knowing about student thinking; it requires gathering information during instruction and then, depending on purpose, interpreting that information in real time for effective use in instruction (Shepard, 2006). The ability to tie assessment evidence to subsequent instruction involves understanding the trajectory of student learning and having instructional strategies that help learners move along that trajectory (Andrade, 2010). Some evidence suggests that research-based frameworks of how students build mathematical understanding can enhance teacher's ability to interpret evidence of student learning and respond productively in light of that evidence (Clements, Sarama, Spitler, Lange, & Wolfe, 2011; Wilson, 2009).

Description of Ongoing Assessment Project

The Ongoing Assessment Project (OGAP) is a mathematics formative assessment program that combines both longstanding ideas of formative assessment and more recent developments in learning trajectories.

OGAP is based on the idea that teachers make more effective instructional decisions that result in improved student learning when they: a) are knowledgeable about how students develop understanding of specific mathematics concepts and the trajectory of student learning of those concepts; b) have tools and strategies that allow them to systematically monitor students' understanding prior to and during instruction; and, c) receive professional development focused on building their knowledge and strategies. OGAP materials have been developed through the distillation of hundreds of studies and articles on mathematics education research on addition, multiplication, and fractions. Organized by content domains, the OGAP formative assessment system includes:

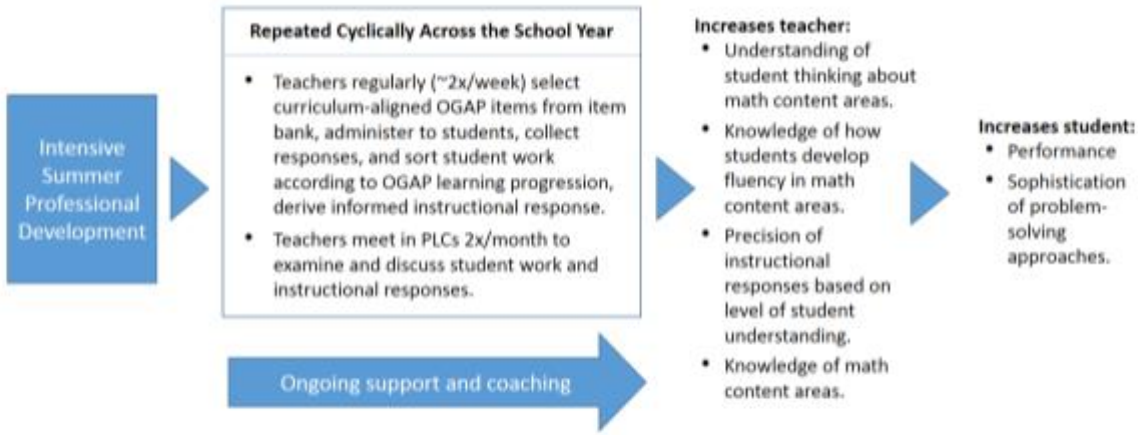
- *Professional development and materials* that focus on specific mathematics topics, develops teachers' knowledge of mathematics and the related research base on student thinking, and provides training in the use of OGAP materials and strategies.
- *Item banks and pre-assessments* designed to elicit students' developing understandings, common errors, and preconceptions or misconceptions. Each item bank has approximately 300 short constructed response items organized by mathematical topics and problem situations (e.g., equal groups, multiplicative change and comparison, measure conversion).
- *Frameworks* that synthesize the problem structures, problem situations, and typical learning progressions for specific mathematics topics to help teachers analyze evidence in student work and make instructional decisions.
- A web-based *eTool* that allows teachers to access items and facilitate teachers' record-keeping and use of formative assessment information in real time.

OGAP Theory of Action

The OGAP theory of action is depicted in Figure 1. Teachers receive multiple days of summer professional development and ongoing support and coaching across the school year. They are asked to regularly select and administer curriculum-aligned open-ended assessment items from OGAP item bank, collect the responses, and sort student work according to the OGAP learning progression, and apply an informed instructional response. To support this work, teachers are expected to meet every other week in professional learning communities to examine and discuss student work and instructional responses.

These regular routines are theorized to increase teacher understanding of student thinking, math content, how student develop conceptual fluency, and to improve the precision of teachers' instructional responses. These increases in teacher capacity are, in turn, intended to lead to improvement in student performance both in terms of accuracy and solution sophistication.

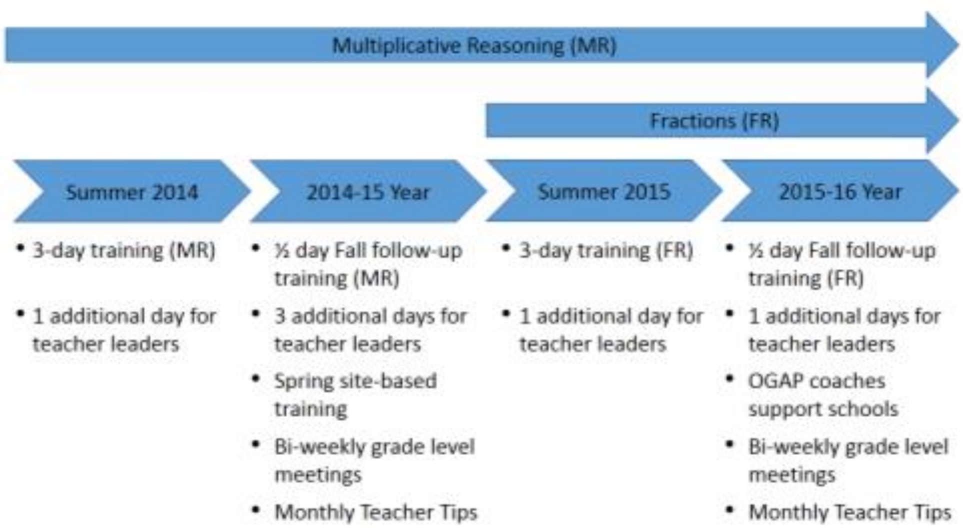
Figure 1. OGAP Theory of Action



OGAP Training and Support in Philadelphia

OGAP provides several components in its training and support for implementing teachers and schools. These are detailed in Figure 2. In year one, teachers begin with multiplicative reasoning, which includes summer professional development and followup training for both teachers and teacher leaders. In year 2, the emphasis of summer training shifts to fractions, with similar additional training during the school year. One difference between year one and year two support was the addition of OGAP coaches in year 2, which OGAP felt would improve attention to OGAP by teachers. A separate paper describes the challenges of supporting implementation across both years.

Figure 2. Summary of OGAP Professional Development and Support in Philadelphia



Randomized Experiment in Philadelphia Area Schools

In the Spring of 2014, CPRE recruited schools in the School District of Philadelphia (SDP) and the neighboring district of Upper Darby (UD) to participate in a randomized experiment of OGAP. Schools were recruited with the promise that they would either receive OGAP training and support for two years (2014-2016) or their school would receive $1,000 for their school activity fund each year. In all, 60 schools agreed to participate, which included 37 schools from the SDP, 13 Philadelphia charter schools, and 10 UD schools. Schools were stratified in each of these categories (SDP, charter, and UD schools) and randomly assigned to the OGAP treatment or control group. Demographics of the schools in the two groups are shown in the sample section below.

School and Student Sample

Our analytic sample focused on a subset of students that met specific selection criteria. Students were excluded if they switched schools, regardless of treatment status, to eliminate any crossover or dosage concerns. In addition, we focused on students who completed the OGAP assessment in each relevant wave, e.g. both the Fall of 2014 and Spring of 2015. Students may not have completed assessments for a number of reasons, including leaving the school during the school year or simply absence during the assessment period. On some occasions teachers or schools did not administer or return the assessments. There was no indication that the distribution of individual demographic information across treatment assignment noticeably changed between the full and analytic sample. For the analytic sample, missing information on student demographics ranged from 4.37% to 8.10% depending on wave and grade. A future revision will use various imputation procedures to handle missing data. Student demographic information was not available for the Fall 2015-Spring 2016 3rd grade cohort. We used 2014-2015 school-level characteristics for both academic years as school characteristics were fairly consistent from year-to-year. This information will be updated when the National Center for Education Statistics releases 2015-2016 information later this year.

The overall analytic sample included in the analysis were 9,081 students in grades 3-5 in 60 schools, 30 treatment and 30 control. Of the treatment schools, 18 were from the SDP, 7 were charter schools in Philadelphia, and 5 were UD schools. In the control group, 19 were from the SDP, 6 were charter schools in Philadelphia, and 5 were UD schools. The samples of schools and students, organized by grade level and treatment condition, is shown in Table 1.

The treatment and control schools were very similar in terms of size and poverty, as measured by percent of students eligible for free/reduced lunch. There were no significant differences at any of the three grades in terms of school size or lunch assistance.

At the student level – and perhaps most importantly – at baseline there were no differences between student performance between the students in the treatment and control schools on the 2014 pre-test in terms of either correctness or sophistication. In fact, on these two measures it remarkable how similar the two groups performed.

There were, however, some differences in the composition of the treatment and control groups across the grade levels. The treatment group had significantly fewer Hispanic students, and more Asian students, at all three grade levels than did the control group. The control group also had significantly more students with disability in the fourth and fifth grades than did students in the treatment group. There were no differences between the treatment and control groups in terms of receipt of free/reduced lunch at any of the three grade levels.

Measures

Students use a variety of strategies to solve multiplicative problems that can be located on a developmental progression. Depending on the strength of students' multiplicative reasoning, they often move back and forth between multiplicative, transitional, additive, and non-multiplicative strategies as they interact with different problem structures and contexts (Kouba & Franklin, 1995). In the OGAP project, teachers learn to use a learning progression framework to continually assess and adapt their instruction to students' developing understanding, aiming to move them towards more sophisticated strategies in a range of multiplicative contexts. For this reason, OGAP puts a premium on students' precision of answer (including correctness and unit labeling) and sophistication of solution response.

In order to assess the intervention's impact on student learning in both correctness and sophistication, the research team developed a measure with three vertically-equated grade-specific forms composed of open-ended items. The assessment asked students to show their work to allow for analysis of their correctness, solution strategies, and errors. The assessment is closely aligned with the CCSSM in terms of the range of problem structures and complexity of number.

To develop the new assessment we first conducted an extensive field test of 30 items constructed to represent the range of problem structures and levels of difficulty expected by the CCSSM at each grade. In the Spring of 2014 we administered the pilot in carefully determined permutations of items (booklets) to a sample of 1400 students from outside of our impact study. We also developed a six level rubric to measure sophistication of strategy in relation to the learning trajectory on each item: (0) non-multiplicative; (1) early additive; (2) additive; (3) early transitional; (4) transitional with models; and (5) multiplicative.

We used an IRT model to analyze the full sample and select 12 items that elicited a range of strategies appropriate to grade level. The final forms were each composed of seven items with four common items. This assessment is an important development not just for measuring outcomes in the OGAP project, but for researchers and practitioners interested in assessing student performance in relation to the more rigorous expectations of the CCSSM.

Analytic Plan

Our two main outcomes are correctness (on a 0-2 scale with 0 incorrect, 1 correct, 2 correct with units labeled) and sophistication (on a 0-5 scale with 0 non-multiplicative; 1 early additive; 2 additive; 3 early transitional; 4 transitional with models; and 5 multiplicative.) scores derived from the OGAP assessments. Students received a baseline assessment of three questions in the fall of 2014 and 2015, and a larger seven question assessment in the spring. The questions were vertically equated across grades, with students in higher grades receiving more advanced assessments. Both the correctness and sophistication scores are an evenly weighted average of the correctness or sophistication scores, respectively, of the individual assessment items. As a sensitivity, an alternative binary correctness score (0 incorrect, 1 correct with or without units labeled) was also assessed and yields similar results.

We treated the data as repeated cross-sections and fit a two-level hierarchical linear model with students nested within schools. The key variable of interest is the treatment indicator for whether the student attended a control (0) or treatment (1) school. The model controls for the student's assessment score in the earlier wave, a vector of student demographics, and a vector of school characteristics. Student demographic information included gender (1 female, 0 male) race/ethnicity (white, black, Hispanic, Asian, other), free/reduced lunch (0 neither, 1 reduced, 2 free), and binary indicators for disability and English language learners. School characteristics included school size, proportion of the student body on free or reduced lunch, charter school status, and the district in which the school was located (Philadelphia or Upper Darby). Standard errors are clustered at the school-level.

We specified three sets of analyses. The first set focused on the Fall 2014 to Spring 2015 time period for 3rd, 4th, and 5th graders to examine the influence of OGAP on student multiplication learning in the first year of the intervention. The second set examined whether any differential gains observed for students in the treatment group persist or expand over the 2015-2016 school year, with 3rd and 4th graders in the Spring of 2015 followed to the Spring of their 4th or 5th grade year, respectively (2016). In addition, a new cohort of 3rd graders are examined from the Fall of 2015 to the Spring of 2016. The third set of analyses examined the same 3rd-4th and 4th-5th cohorts as in the second set, but used the Fall of 2014 as the pre-score to determine whether any gains for students in the treatment group persist across both years of the intervention.

Results

The results for correctness and sophistication are shown in tables 2 and 3 respectively. Each table has three sets of results. The first set of results in each table is for 3rd, 4th and 5th grade Spring 2015 performance controlling for Fall 2014 performance, as well as for student and school demographics and appropriately nesting students within schools. These can be seen as the first year OGAP results assessing student multiplicative reasoning.

The second set of results in each table shows Spring 2016 performance, controlling for Spring 2015 performance[2], as well as for student and school demographics and appropriately nesting students within schools. These can be seen as the second year OGAP impact results assessing student multiplicative reasoning.

The third set of show Spring 2016 performance, controlling for Fall 2014 performance, as well as for student and school demographics and appropriately nesting students within schools. These can be seen as the net effect of two years of OGAP treatment.[3]

Focusing on the treatment effects for correctness in Table 2, we can see that in the first year of OGAP (2014-2015) that the students in the treatment schools statistically outperformed the students in the control schools in grades 3, 4 and 5, after controlling for prior performance and student and school demographics. In the second year, although OGAP students performed, on average, higher than did students in the control schools, the differences were not statistically significant in any of the three grades, after controlling for prior performance and student and school demographics. The 2014-2016 treatment results in the final two columns of Table 2 indicate that students in the treatment group in grades 4 and 5 outperformed those in the control group controlling for baseline performance. This shows that while most of the gains in student performance occurred during year one of their OGAP experience, these gains persisted over their second year in an OGAP school.

Examining the effects for sophistication, shown in Table 3, we can see a similar story about the impacts of OGAP on student solution sophistication. After learning from teachers with one year

[2] Except for 3rd grade, who were administered a pre-test in the Fall of the 2015-2016 school year.

[3] There are no 3rd grade two year effects because there was no baseline 2014 assessment for second graders who became third grades in the 2015-16 school year.

of OGAP professional development, the students in all three grades in the treatment schools significantly outperformed those in the control schools, after adjusting for prior achievement and student and school characteristics. These effects did not sustain into year two (2015-16), where, on average, there were no significant differences between the performance of students in the treatment schools and students in the control schools after controlling for prior achievement and student and school characteristics. For solution sophistication, the two year effects show that 4th and 5th grade students in 2016 in the treatment schools maintained their higher levels of performance compared to their peers in the control schools.

Note to Hilda and Tonya: We plan to add some result information about differences between students and school types to the AERA version, but we want to get this to you.

Discussion

The results show strong effects in multiplicative reasoning outcomes in year one of the OGAP intervention, but not in year two. We have several hypotheses that might explain these effects. First, it is possible that as OGAP shifted emphasis from multiplication to fractions, the lessening emphasis on fractions reduced the attention to multiplication. Second, when observing implementation support in year 2, we noted multiple challenges in having teachers meet regularly in PLCs and to use OGAP items on a consistent basis in year 2. We have previously seen low implementation levels in our data, and this may be reducing student exposure to OGAP principles. Third, particularly in the School District of Philadelphia, there is substantial teacher and student mobility. This raises interesting questions about the variations in teacher and student exposure to OGAP, which are not yet part of our modeling but suggest an important next step for analysis.

Table 1. Descriptive Student and School Characteristics, 2014-2015

	3rd Grade			4th Grade			5th Grade		
School Demographics	Control	Treatment		Control	Treatment		Control	Treatment	
School Size (hundreds)	5.94	5.67		5.94	5.67		5.90	5.67	
	(1.88)	(2.25)		(1.88)	(2.25)		(1.93)	(2.25)	
Free Lunch Percent	81.42	84.12		81.42	84.12		80.13	84.12	
	(21.28)	(15.89)		(21.28)	(15.89)		(21.45)	(15.89)	
Student Characteristics									
Female	51.23	49.55		48.21	49.63		51.88	51.89	
Race	--	--		--	--		--	--	
White	18.10	20.88	+	18.24	18.19		18.64	18.46	
Black	48.00	49.17		48.21	50.44		49.01	50.11	
Hispanic	20.89	10.97	***	18.43	12.29	***	17.75	11.97	***
Asian	7.60	15.13	***	7.84	14.60	***	9.25	16.39	***
Other	5.41	3.86	*	7.27	4.48	***	5.35	3.06	**
Lunch	--	--		--	--		--	--	
None	31.98	31.32		29.47	31.57		33.38	34.14	
Reduced	1.59	1.97		1.19	1.63		1.99	1.64	
Free	66.43	66.72		69.34	66.80		64.63	64.22	
Disability	13.45	12.63		16.87	13.31	**	15.22	12.26	*
ESL	11.97	15.05	*	8.71	10.25		8.91	9.34	
Fall 2014 Correctness	0.51	0.50		0.38	0.37		0.62	0.63	
	(.51)	(.50)		(.47)	(.42)		(.57)	(.55)	
Fall 2014 Sophistication	0.65	0.63		1.08	1.13		1.91	1.95	
	(.74)	(.69)		(1.08)	(1.10)		(1.48)	(1.45)	
N	1,829	1,322		1,595	1,473		1,459	1,403	

Table 2. Multivariate Regression Model Predicting OGAP Correctness Score[a]

	2014-2015			2015-2016			2014-2016	
	3rd Grade	4th Grade	5th Grade	3rd Grade	4th Grade	5th Grade	4th Grade	5th Grade
Fall 2014 Score	0.333 ***	0.429 ***	0.372 ***				0.289 ***	0.405 ***
	(0.018)	(0.020)	(0.017)				(0.022)	(0.023)
2015 Score[b]				0.394 ***	0.471 ***	0.579 ***		
				(0.014)	(0.026)	(0.029)		
Treatment	0.114 *	0.116 *	0.108 **	0.044	0.035	0.017	0.111 ***	0.085 +
	(0.054)	(0.041)	(0.036)	(0.039)	(0.033)	(0.048)	(0.037)	(0.046)
Student Characteristics		--	--	--	--	--	--	--
Female	0.023	-0.003	0.015	N/A	-0.006	-0.004	-0.005	-0.009
	(0.014)	(0.016)	(0.014)		(0.016)	(0.016)	(0.018)	(0.017)
Race	--	--	--	--		--	--	--
Black	-0.200 ***	-0.120 ***	-0.114 ***	N/A	-0.071	-0.091 *	-0.135	-0.155 ***
	(0.034)	(0.028)	(0.028)		(0.043)	(0.036)	(0.055)	(0.038)
Hispanic	-0.115 ***	-0.049	-0.066 *	N/A	-0.007	-0.056	-0.055	-0.052
	(0.042)	(0.039)	(0.030)		(0.044)	(0.041)	(0.060)	(0.050)
Asian	0.120 *	0.146 ***	0.207 ***	N/A	0.204 ***	0.147 **	0.283 ***	0.196 ***
	(0.057)	(0.031)	(0.031)		(0.050)	(0.048)	(0.068)	(0.049)
Other	-0.059	-0.042	0.005	N/A	-0.040	-0.030	-0.049	-0.025
	(0.049)	(0.043)	(0.041)		(0.037)	(0.039)	(0.048)	(0.044)
Lunch	--	--	--	--		--	--	--
Reduced	0.043	0.003	-0.008	N/A	-0.047	-0.018	-0.035	0.027
	(0.055)	(0.039)	(0.058)		(0.057)	(0.058)	(0.059)	(0.058)
Free	0.017	-0.074 ***	-0.060 **		-0.053 ***	-0.030	-0.037	-0.072 **

	(0.019)	(0.015)	(0.019)	N/A	(0.016)	(0.021)	(0.020)	(0.025)
Disability	-0.212 ***	-0.212 ***	-0.160 ***		-0.137 ***	-0.067 *	-0.207 ***	-0.126 ***
	(0.024)	(0.019)	(0.021)		(0.020)	(0.028)	(0.022)	(0.029)
ESL	-0.094 *	-0.149 ***	-0.154 ***	N/A	-0.051 **	-0.092 **	-0.058 **	-0.161 ***
	(0.039)	(0.026)	(0.029)		(0.020)	(0.031)	(0.022)	(0.029)
School Demographics		--	--	--	--	--	--	--
School Size (hundreds)	-0.012	-0.006	-0.007	-0.007	0.005	0.002	-0.006	-0.005
	(0.014)	(0.011)	(0.007)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)
Free Lunch Percent	-0.238 *	-0.285 ***	-0.257 ***	-0.398 ***	-0.164 *	-0.263 *	-0.239 *	-0.293 *
	(0.108)	(0.107)	(0.082)	(0.133)	(0.071)	(0.122)	(0.097)	(0.119)
Charter	0.029	-0.137 ***	-0.098 *	0.003	0.015	0.000	0.109 *	-0.056
	(0.061)	(0.038)	(0.038)	(0.060)	(0.048)	(0.048)	(0.055)	(0.054)
Upper Darby	-0.074	-0.094	-0.029	-0.067	0.062 +	0.044	0.062	0.028
	(0.066)	(0.062)	(0.043)	(0.059)	(0.035)	(0.057)	(0.042)	(0.052)
Constant	0.783 ***	0.784 ***	0.641 ***	0.646 ***	0.392 ***	0.491 ***	0.648 ***	0.774 ***
	(0.154)	(0.137)	(0.092)	(0.135)	(0.084)	(0.120)	(0.109)	(0.129)
N	3,160	3,086	2,868	3,389	2,541	2,364	2,288	2,266

+ p≤.10, * p≤ .05, ** p≤.01, *** p≤.001

[a] The analytic sample is restricted to students not missing information on the dependent and independent variables. All models are specified as repeated measures with children nested within schools. Standard errors are clustered at the school level.

[b] Fall 2015 for 3rd grade, Spring 2015 for 4th and 5th grade

Table 3. Multivariate Regression Model Predicting OGAP Sophistication Score[a]

	2014-2015			2015-2016			2014-2016	
	3rd Grade	4th Grade	5th Grade	3rd Grade	4th Grade	5th Grade	4th Grade	5th Grade
Fall 2014 Score	0.468***	0.497***	0.453***				0.538***	0.458***
	(0.033)	(0.026)	(0.023)				(0.044)	(0.026)
2015 Score[b]				0.501***	0.651***	0.639***		
				(0.029)	(0.031)	(0.020)		
Treatment	0.162+	0.347**	0.260**	0.025	0.172	-0.020	0.308*	0.260+
	(0.097)	(0.123)	(0.093)	(0.091)	(0.110)	(0.146)	(0.129)	(0.141)
Student Characteristics		--	--	--	--	--	--	--
Female	-0.151***	-0.106**	-0.036	N/A	-0.031	-0.033	-0.119*	-0.096*
	(0.027)	(0.040)	(0.041)		(0.046)	(0.044)	(0.054)	(0.048)
Race	--	--	--	--	--	--	--	--
Black	-0.341***	-0.317***	-0.368***	N/A	-0.232+	-0.277**	-0.417**	-0.407***
	(0.064)	(0.094)	(0.083)		(0.122)	(0.098)	(0.142)	(0.098)
Hispanic	-0.112+	-0.234+	-0.194+	N/A	-0.004	-0.214+	-0.095	-0.253*
	(0.061)	(0.121)	(0.102)		(0.139)	(0.118)	(0.167)	(0.123)
Asian	0.366***	0.348**	0.399***	N/A	0.452***	0.306**	0.711***	0.472***
	(0.107)	(0.120)	(0.101)		(0.137)	(0.106)	(0.170)	(0.125)
Other	-0.112	-0.252*	0.086	N/A	-0.279*	-0.099	-0.308*	-0.190
	(0.096)	(0.124)	(0.126)		(0.114)	(0.123)	(0.143)	(0.123)
Lunch	--	--	--	--	--	--	--	--
Reduced	0.061	-0.074	-0.149	N/A	-0.128	-0.019	-0.149	-0.072
	(0.105)	(0.112)	(0.141)		(0.149)	(0.102)	(0.169)	(0.134)
Free	-0.004	-0.225***	-0.172**	N/A	-0.125**	0.022	-0.092	-0.129+
	(0.031)	(0.053)	(0.056)		(0.046)	(0.051)	(0.057)	(0.068)

Disability	-0.390***	-0.671***	-0.682***	N/A	-0.412***	-0.275**	-0.622***	-0.548***
	(0.056)	(0.055)	(0.064)		(0.085)	(0.092)	(0.097)	(0.106)
ESL	-0.246***	-0.402***	-0.359***	N/A	-0.192**	-0.149+	-0.265**	-0.393***
	(0.069)	(0.088)	(0.079)		(0.069)	(0.085)	(0.097)	(0.087)
School Demographics		--	--	--	--	--	--	--
School Size (hundreds)	-0.029	-0.020	-0.025	-0.037+	-0.01	0.013	-0.046	-0.004
	(0.022)	(0.036)	(0.020)	(0.021)	-0.031	(0.027)	(0.034)	(0.028)
Free Lunch Percent	-0.585**	-0.836***	-0.587*	-0.741**	-0.561*	-0.740*	-0.859**	-1.127**
	(0.217)	(0.316)	(0.266)	(0.276)	-0.262	(0.368)	(0.280)	(0.383)
Charter	0.104	-0.410**	-0.270*	0.081	-0.072	-0.156	0.097	-0.283+
	(0.099)	(0.150)	(0.133)	(0.154)	-0.122	(0.158)	(0.141)	(0.161)
Upper Darby	-0.271**	-0.227	-0.002	-0.239*	0.161	0.073	0.003	0.078
	(0.104)	(0.148)	(0.118)	(0.120)	-0.103	(0.178)	(0.130)	(0.161)
Constant	1.962***	2.575***	2.332***	1.710***	1.625***	1.720***	2.704***	2.963***
	(0.234)	(0.352)	(0.303)	(0.314)	-0.303	(0.326)	(0.349)	(0.392)
N	3,153	3,068	2,863	3,389	2,540	2,360	2,287	2,263

+ p≤.10, * p≤ .05, ** p≤.01, *** p≤.001

[a] The analytic sample is restricted to students not missing information on the dependent and independent variables. All models are specified as repeated measures with children nested within schools. Standard errors are clustered at the school level.

[b] Fall 2015 for 3rd grade, Spring 2015 for 4th and 5th grade