YEAR-ONE IMPACT AND PROCESS RESULTS FROM THE

i3 IMPLEMENTATION OF *TEACH TO ONE: MATH*

Douglas D. Ready

Katherine Conn

Elizabeth Park

David Nitkin

November 28, 2016

CPRE

CONSORTIUM FOR POLICY RESEARCH IN EDUCATION

TEACHERS COLLEGE

COLUMBIA UNIVERSITY

_____

<center>CONTENTS</center>

New Classrooms Innovation Partners was awarded an Investing in Innovation (i3) grant from the U.S. Department of Education to expand its *Teach to One: Math* model in grades 5-8 in three schools and grades 6-8 in two schools in Elizabeth, NJ beginning in September, 2015, and continuing through June, 2018. The Consortium for Policy Research in Education (CPRE) at Teachers College, Columbia University, is engaged in a four-year evaluation of these efforts. The evaluation has two primary strands: 1) a methodologically rigorous quantitative study of TtO's causal impact on student mathematics performance, and; 2) a deep qualitative analysis of TtO implementation processes, combined with staff interviews and classroom observations in each of the five TtO schools. This report describes the results of both strands for the first year of implementation, the 2015-16 academic year.

Quantitative Impact Evaluation: Year One

Our primary causal analyses entailed a Comparative Interrupted Time Series (CITS) approach (also known as the difference-in-differences technique), with a sample of 31,734 student-level measurements, nested within 143 cohorts, located within five TtO schools and a comparable group of 13 non-TtO schools, all in Elizabeth, NJ. We used this technique to compare changes in state-mandated standardized test scores among TtO schools before and after program implementation to the changes in outcomes among a similar group of schools in Elizabeth that did not implement TtO during the same period. Results suggest that:

- Overall, across all grades, we found no significant impact of TtO on student mathematics performance as measured by state-mandated assessments ($p>.10$). However, we found a moderate, marginally significant negative effect for fifth grade in the three schools implementing TtO in that grade (*effect size* [*ES*] = -0.371 *SD*s; $p<.10$).

Because TtO matches students to appropriate mathematics content regardless of student grade level, we were concerned that academic gains made among initially low-achieving students might be undetected by traditional state assessments, which typically include a much narrower band of grade-level content. Fortunately, students in the district were also administered the Measures of Academic Progress (MAP) assessments in the fall, winter, and spring of the 2015-16 academic year. We used these student-level MAP scores as outcomes to explore growth during three periods: fall to winter; winter to spring; and fall to spring. Results suggest that:

- TtO students gained modestly fewer mathematics skills between the fall and spring MAP assessments compared to students in comparison schools (*ES* = -0.283; $p<.05$).
- However, this negative effect occurred solely during the initial fall-to-winter implementation period. TtO and non- TtO students gained mathematics skills at comparable rates during the subsequent winter-to-spring period ($p>.10$).

Process Evaluation: Year One

CPRE Researchers visited all five TtO schools, observed classes, and interviewed a total of 44 participants, including five school principals, two vice-principals, five Math Directors, 16 teachers (Math Advisors), five Teacher Residents (TRs) for special education students, five Teacher Residents (TRs) for bilingual students, and six New Classrooms staff members (two Site Operations Managers (SOMs), two Operations and Technology Associates (OTAs), and two coaches. Interviews were semi-structured and lasted approximately 30-45 minutes each.

Overall impressions. Almost universally, teachers stated that they understood and endorsed the TtO "theory of action," including differentiation of content, instantaneous feedback through exit slips, and diversity of learning modalities. Teachers particularly enjoyed how TtO leveraged technology to allow them to work with small groups of students in particular modalities. Teachers also asserted that TtO increased student agency, and that it provided many students with the resources to thrive should they choose to take advantage of them. However, teachers also claimed that students' intrinsic motivation to maximize TtO's potential varied, and that TtO's extrinsic rewards may not have been equally valued by all students. Although teachers were virtually unanimous in their support of TtO's philosophy, many were also quite concerned that by not necessarily exposing all students to grade-level content, student performance on state grade-level assessments would suffer, as would their own performance ratings within the State's teacher accountability system.

TtO modalities and processes. Some teachers expressed concern that particular modalities provided fewer opportunities for critical thinking and applications to real-world problems, although these teachers also claimed that the Task modality partially addressed these gaps. Similarly, teachers missed the opportunity to review and provide feedback on student work. Many felt that multiple-choice-only exit slips made it difficult to understand students' thinking; some teachers sought more writing opportunities for students to explain their understanding. Differentiated homework also made it difficult for teachers to provide feedback to students. Interestingly, several teachers asserted that although TtO personalized *content*, it did not assist in the personalization of teacher/student relationships, which these teachers claimed were more difficult to establish and maintain, due largely to the shorter class sessions, and the fact that they interacted with larger numbers of different students each day.

Views of TtO logistics. Teachers were generally satisfied with New Classrooms' support and responsiveness, including the placement of staff on-site and the willingness to customize the program to meet school needs. Teachers found the TtO portal to be useful, and particularly appreciated the immediacy and accessibility of student data, although several schools faced sporadic technical problems with student log-ins and bandwidth barriers, particularly during exit slip sessions. Teachers also expressed a desire to access multiple student profiles simultaneously to better understand how students were performing as a whole. More pressing were concerns that TtO included insufficient content for English Language Learners and Students with Disabilities. The open classroom layout, which many teachers asserted increased noise and distractions, also presented challenges for some teachers and schools, as did the shorter instructional sessions.

Conclusions

Unlike many educational reforms, TtO represents a fundamental disruption of traditional classroom processes, and provides a distinctive approach to mathematics teaching and learning. In short, it asks a great deal of teachers and students. Given this, we were both surprised and impressed with teachers' willingness and openness to the model, particularly with teachers who had been in the classroom, in some cases, for decades. Among virtually all teachers we interviewed, support for the logic behind TtO was broad. Many of the concerns they did raise about particular elements of TtO are well-known by New Classrooms staff, who have, according to these teachers, responded appropriately where possible. Other criticisms of the model may take further discussion and negotiation to fully resolve.

The quantitative impact results suggest that TtO teachers and students may have experienced an initial adjustment period in which student learning suffered. However, there is

some evidence that as TtO schools became more familiar with the program during the latter half of the initial implementation year, student learning stabilized. Our evaluation of TtO in years two and three will establish whether outcomes among TtO students continued on this upward trajectory. We look forward to visiting with these same teachers and schools again during the second year of TtO implementation in the Elizabeth Public Schools.

Practitioners and researchers have long explored the appropriate way to manage academic diversity in the classroom (Corno, 2008). Because students in the U.S. are typically grouped into same-age classrooms regardless of academic background, teachers confront a tremendous amount of variability in student skills. Indeed, despite considerable stratification and segregation across schools and classrooms, almost 62% of variability in fifth-grade mathematics ability remains within classrooms, with the additional variability split relatively evenly between classrooms in the same schools and between schools (Martinez, Stecher, & Borko, 2009). Such academic diversity has historically been handled via ability grouping in the lower grades (Pallas, et al., 1994; Barr & Dreeben, 1983) and curricular differentiation and "tracking" in the upper grades (Lee & Ready, 2009; Oakes, 1985). Unlike these traditional approaches, which typically do not serve low-achieving or very high-achieving students well, advocates have instead argued for "adaptive instruction" (Snow,1980) and "personalized learning" approaches (Gates Foundation, 2014), which seek to treat students as individuals and respond to their needs amidst the diversity of the collective social classroom environment (Corno, 2008). Similar to Vygotsky's (1978) "zone of proximal development," personalized approaches aim to situate students in cognitive spaces that are slightly beyond their current skills, and "scaffold" their learning by concentrating efforts into more manageable sets of tasks (Bruner, 1978).

Although various efforts to implement personalized instruction have existed for decades, technological innovations over the past several years have made it much easier to individualize instruction based on each student's level of content mastery and developmental trajectory (Wolf, 2010). These contemporary approaches entail individualized and personalized learning plans based on student-level data, recognize progress that is based on demonstrated knowledge rather than seat time, and employ multiple and flexible pedagogical and learning environments (Gates Foundation, 2014). With the assistance of "blended learning" approaches, which combine computer-based and live teacher-directed instruction, several Charter Management Organizations have built models that focus on personalized learning approaches, including Summit Public Schools and Rocketship Education, as have select private schools, such as AltSchool.

One such program, which involves both personalization and blended learning, is *Teach to One: Math* (TtO), developed by New Classrooms Innovation Partners. TtO currently serves roughly 13,000 fifth to eighth grade students in 40 charter and regular public schools in 10 states

and Washington D.C.[1] The model simultaneously integrates multiple technology-enabled and live instruction modalities, including teacher-led instruction; small-group and peer-to-peer collaborations; independent computer-based practice; coached virtual instruction; virtual reinforcement; and project-based task sessions that extend over multiple days. To support these modalities, classroom space is dramatically transformed, with one large space supporting several (often multi-grade) classes learning simultaneously across multiple work spaces.

After extensive baseline assessments, each student is provided with a personalized curriculum at the start of the year. Then, based on the results of short daily assessments in combination with curricular content maps, an algorithm generates a unique daily learning plan for each student according to his or her learning needs. Key to the TtO model is that students' daily work, in terms of both content and learning modality, is determined by their level of skill mastery. Teacher and student portals provide real-time information on student progress and enable students and teachers to navigate their schedules and lessons and review performance history. Fundamental to the model is the belief that student progress in mathematics is limited if students lack prerequisite mathematics skills. To this end, the model permits students to engage material that may be several grade levels below their current grade, drawn from a 15,000-item lesson bank.

TtO has been highlighted in the New York Times, Wall Street Journal, the Atlantic, and Time Magazine, which named it one of the top 50 "inventions of 2009." At a recent conference Bill Gates referred to TtO as "the future of math" (Newcomb, 2016). Despite these adulations, TtO has not been subject to a recent evaluation that could establish its causal impact, and it remains unclear whether TtO improves student outcomes.[2] New Classrooms was awarded an Investing in Innovation (i3) grant from the U.S. Department of Education to expand the TtO model into five schools in Elizabeth, New Jersey, beginning in September, 2015, and continuing through June, 2018. Using both quantitative and qualitative methodological approaches, this report explores the causal effect of TtO on student mathematics achievement in year one of the i3 implementation, and describes how teachers and staff in the TtO schools interpreted the model's efficacy and impact.

---

[1] www.newclassrooms.org

[2] Rigorous, but small-scale analyses of a previous version of Teach to One in New York City, then known as School of One, reported mixed results (see Cole, Kemple, & Segeritz, 2012; Rockoff, 2015).

QUANTITATIVE IMPACT EVALUATION

Data and Methods

To measure the impact of the TtO program on student learning in mathematics, we use data from five treatment and 13 comparison schools in Elizabeth, New Jersey, a high-minority, high-poverty school district within the New York City metropolitan area. In Elizabeth, 26 K-8 schools and six high schools serve approximately 23,000 district students, of whom 90% are either black or Hispanic, with a similar proportion eligible for free or reduced-price lunch. Over 13% of students are enrolled in bilingual classes. The fact that both TtO and comparison schools are drawn from the same relatively homogenous district supports our ability to make claims about TtO's impact on student achievement. To further ensure comparability, we eliminated control schools without substantial overlap in terms of baseline achievement and student demographics. These schools included two district-wide gifted and talented schools, and three additional relatively higher-achieving schools. We also eliminated two schools that were implementing TtO, but were not involved in the i3 implementation, and had been engaged with TtO for several years prior. Two of the five TtO treatment schools implemented the program in only grades six through eight. For these schools, fifth graders are eliminated from the sample for both pre-treatment and treatment years.[3] Our final analytic sample includes 31,734 student-level measurements, nested within 143 student cohorts, located within five TtO and 13 non-TtO schools, all in Elizabeth, NJ.

Analytic Approaches

The quantitative year-one analyses relied on two distinct approaches and sets of outcomes to provide early indication of the extent to which TtO improves student learning in mathematics. Our primary causal analyses entailed a comparative interrupted time series (difference-in-differences) approach using state-mandated standardized assessment scores as outcomes. Our secondary analyses involved hierarchical linear gain-score models that explored student learning as captured by the Measures of Academic Progress (MAP) during the initial implementation year (2015-16).

---

[3] The relatively small number of potential treatment and control schools would not support traditional school-level propensity score matching.

Comparative interrupted time series models. To obtain causal estimates of the impact of TtO on student mathematics learning, we used a comparative interrupted time series approach (CITS; Bloom, 2003; Shadish, Cook, & Campbell, 2002), constructed within a multi-level framework (see Raudenbush & Bryk, 2002), with students nested within school cohorts, nested within schools. Within the context of the current evaluation, the CITS approach compares changes in outcomes among TtO schools before and after program implementation to the changes in outcomes among a similar group of schools that did not implement TtO in the same district during the same period. Below we describe a multi-level baseline means model, which reduces sensitivity to noise in student achievement and reduces the risk of mis-specified slopes (see Bloom, 2003).

Level-1 Model (students): $\quad Y_{ijk} = \pi_{0jk} + \pi_{1jk}(X_{ij} - \overline{X_j}) + \ldots + e_{ijk}$

Level-2 Model (cohorts): $\quad \pi_{0jk} = \beta_{00k} + \beta_{01k}(TrtYr1_{jk} - \overline{X_k}) + \beta_{02k}(X_{jk} - \overline{X_k}) + \ldots + r_{0jk}$

Level-3 Model (schools): $\quad \beta_{00k} = \gamma_{000} + \gamma_{001}(TtO_k) + \gamma_{002}(X_k) + \ldots + u_{00k}$
$\qquad\qquad\qquad\qquad\quad \beta_{01k} = \gamma_{010} + \gamma_{011}(TtO_k) + u_{01k}$

where: Y_{ijk} = test score for student i in cohort j in school k
$\qquad X_{ij}$ = student covariates, centered around the within-cohort mean
$\qquad e_{ijk}$ = the error term associated with child ijk, assumed to be normally distributed with a mean of zero and a constant Level-1 variance, σ^2
$\qquad TrtYr1_{jk}$ = 1 if treatment year 1 (2015-2016 academic year), 0 if pre-treatment year (2008/09-2014/15 academic years)
$\qquad X_{jk}$ = cohort (within-school) covariates (aggregates of student-level measures), centered around the within-school mean
$\qquad r_{0jk}$ = error associated with cohort j in school k
$\qquad \gamma_{000}$ = the mean for comparison (non-TtO) schools during pretreatment years
$\qquad \gamma_{001}$ = the average difference between TtO and non-TtO schools for pretreatment years
$\qquad X_k$ = school-level covariates (aggregates of student-level measures), grand-mean centered
$\qquad u_{00k}$ = error associated with school k
$\qquad \gamma_{010}$ = the average difference between pretreatment and year 1 for the comparison schools
$\qquad \gamma_{011}$ = *our focus: the treatment effect for year 1 (the difference-in-difference estimator).*

Centering the student- and cohort-level predictors around the group means—referred to as adaptive centering in multilevel contexts—produces estimates that are directly analogous to fixed effects models, but with several advantages (Raudenbush, 2009). An important result—one required of the CITS approach—is that students are compared to students in the same cohort

attending the same school, and cohort comparisons are to other same-school cohorts. All models were run using HLM 7.01 (Raudenbush, Bryk, & Congdon, 2013).

Hierarchical linear gain-score models. One of the signature elements of the TtO model is the matching of students to appropriate content regardless of student grade level. For example, a seventh-grade TtO student may initially encounter skills that are typically introduced in third grade. This same student may progress to fifth-grade material by the end of the initial TtO year—a notable achievement. However, such growth at the lower-end of the achievement distribution may not be detected on traditional state assessments, which typically target a much narrower band of grade-level content. Fortunately, in addition to the state-mandated standardized assessments, students in all Elizabeth Public Schools were also administered the MAP assessments in both Mathematics and ELA, in the fall, winter, and spring of the 2015-2016 academic year. We use student-level MAP scores as outcomes in simple two-level hierarchical linear models (HLMs). The aim here is to shed further light on student mathematics learning in TtO and non-TtO schools during the first year of implementation, not to provide causal estimates of TtO's impact.[4] In particular, we sought to explore growth during the initial implementation period (fall to winter), the subsequent semester (winter to spring), and the full academic year (fall to spring). Our thought was that growth may have suffered initially as students and teachers became familiar with the program, but improved in the later half of the academic year. Again, unlike the CITS models described above, which afford causal estimates of TtO's impact, the simple HLM models are merely suggestive, but hopefully provide more finely tuned evidence not supplied by the CITS models. We describe these models below as:

Level-1 Model (students): $Y_{ij} = \beta_{0j} + \beta_{1j}(FALL\ ACH_{ij} - \overline{X}_j) + \beta_{.j}(X_{ij} - \overline{X}) + r_{ij}$

Level-2 Model (schools): $\beta_{0j} = \gamma_{00} + \gamma_{01}(TtO_j) + \gamma_{0.}(W_j) + u_{0j}$
                                                 $\beta_{1j} = \gamma_{10} + \gamma_{11}(TtO_j) + u_{1j}$

where: Y_{ij} = predicted mathematics gain for student i in school j, adjusted for fall mathematics
            achievement (group-mean centered) and a vector of student-level characteristics (X_{ij})
            centered around their respective grand means;
         r_{ij} = the error term associated with child ij, assumed to be normally distributed with a

---

[4] The main impediment is that EPS did not administer the MAP assessments in Fall, 2014. Hence, we cannot employ the CITS approach to compare fall-to-spring growth during the 2014-15 to fall-to-spring growth during the implementation year (2015-16).

mean of zero and a constant Level-1 variance, σ^2;

TtO = indicates school implemented *Teach to One* (1=yes, 0=no)

W_j = a vector of school-level covariates (aggregates of the student-level measures);

u_{0j} = error associated with school *j*;

γ_{11} = estimate of how the TtO effect on gain varies as a function of student initial ability (the "slope-as-outcome" approach);

u_{1j} = error term for the fall achievement/math gain slope associated with school *j*.

Standardized outcomes. The CITS model described above employs seven years of prior student-level achievement and socio-demographic data. As in many states, New Jersey's school accountability system has experienced considerable flux over the past several years, with changing standardized assessments over time. The models presented here employ data from two large-scale mathematics assessments. For pre-test academic years 2008-2009 through 2013-2014, we use student-level results on the New Jersey Assessment of Skills and Knowledge (NJASK).[5] Beginning with the 2014-15 academic year, New Jersey adopted assessments organized by the Partnership for Assessment of Readiness for College and Career (PARCC).[6] For pre-test year 2014-15, as well the year-one post-test scores (2015-16), we use the student-level PARCC scale scores. All test scores are z-scored (standardized) within grade and year. The second set of analyses use student-level data on the Measures of Academic Progress (MAP) mathematics assessment, created and managed by the Northwest Evaluation Association (NWEA).[7] MAP scores are standardized within grades.

Student-level measures. Student-level covariates for both the CITS and HLM gain-score models include gender (female=1, male=0), limited English proficiency and special education status (yes=1, no=0), and separate indicators of race/ethnicity (black, Hispanic and Asian/Pacific Islander students, all compared to whites). We also incorporate separate measures of free- and reduced-price lunch status (yes=0, no=1) and grade level (fifth, sixth, and eighth, compared to seventh). The cohort-level (Level-2) models for the CITS analyses, and the school-level models for both the CITS (Level-3) and the HLM gain-score models (Level-2) include aggregate indicators of all student-level measures except grade.

---

[5] For more information on NJASK, see www.nj.gov/education/assessment/es/njask/

[6] For more information on PARCC, see www.parcconline.org/about

[7] For more information on MAP, see www.nwea.org

Impact Results

Table 1 provides descriptive information on school socio-demographic and academic characteristics across TtO and non-TtO comparison schools. Although the CITS models employ covariate adjustments, it is important that the treatment and control schools are as similar as possible on all pre-treatment indicators. This is indeed what we find here, with no statistically significant differences in student demographics between TtO and non-TtO schools. Both sets of schools have somewhat larger proportions of male students, and have LEP and special education enrollments that are roughly 11-14% of their student populations. These schools also serve high-poverty student clientele, with between 80-90% of students eligible for free- or reduced-price lunch, and a largely non-white student demographic, with enrollments roughly 90% black and Hispanic.

In terms of standardized mathematics performance, the TtO and non-TtO comparison schools have virtually identical average baseline mathematics achievement test scores. To further establish baseline equivalence, we constructed a reduced (unadjusted) form of the multilevel CITS model described above that employed only pre-implementation years as outcomes and the TtO school-level indicator as the sole predictor. The TtO estimate was virtually zero (*ES*=0.013) and non-significant (*p*=0.893), suggesting no relationship between TtO participation and prior achievement. Note, however, that after the first year of TtO implementation, TtO schools scored roughly one-quarter standard deviation below the comparison schools (*p*<.001). This provides some early (though descriptive) evidence that students in TtO schools may have achieved at lower levels during the first year of TtO i3 implementation in Elizabeth.

Table 1. Academic and Socio-Demographic Characteristics for TtO and non-TtO Schools

	TtO Schools (*n*=5)	Non-TtO Schools (*n*=13)
% Female	47.2	48.3
% Limited English Prof.	13.8	11.3
% Special Education	14.3	10.8
% Free Lunch	73.6	80.4
% Reduced-Price Lunch	9.0	8.2
% Asian	1.5	2.3
% Black	24.0	26.2
% Hispanic	63.9	65.4
% White	11.6	9.1
Baseline Math Ach.[1]	-0.022	0.007
Year-1 Math Ach.***	-0.175	0.059

***p<.001. Other than year-1 math achievement, no differences significant at *p*<.10.
[1] test scores standardized (z-scored) within grades and years.

Figure 1 below is a pictorial display of the CITS model. Note first that over the past decade, achievement in TtO and non-TtO schools has been quite constant, fluctuating only slightly within a very narrow range of achievement. The dotted line represents the mean performance of TtO schools prior to TtO implementation projected into the initial implementation year (2015-16). Note that the actual TtO school mean for the implementation year is below what was predicted based on TtO school prior performance. In contrast, as indicated by the dashed line, non-TtO school performance is very slightly above what prior performance had predicted. The CITS model compares the difference in scores among TtO schools pre- and post-implementation, to the difference in scores among non-TtO schools pre- and post-implementation; hence, the "difference-in-differences."

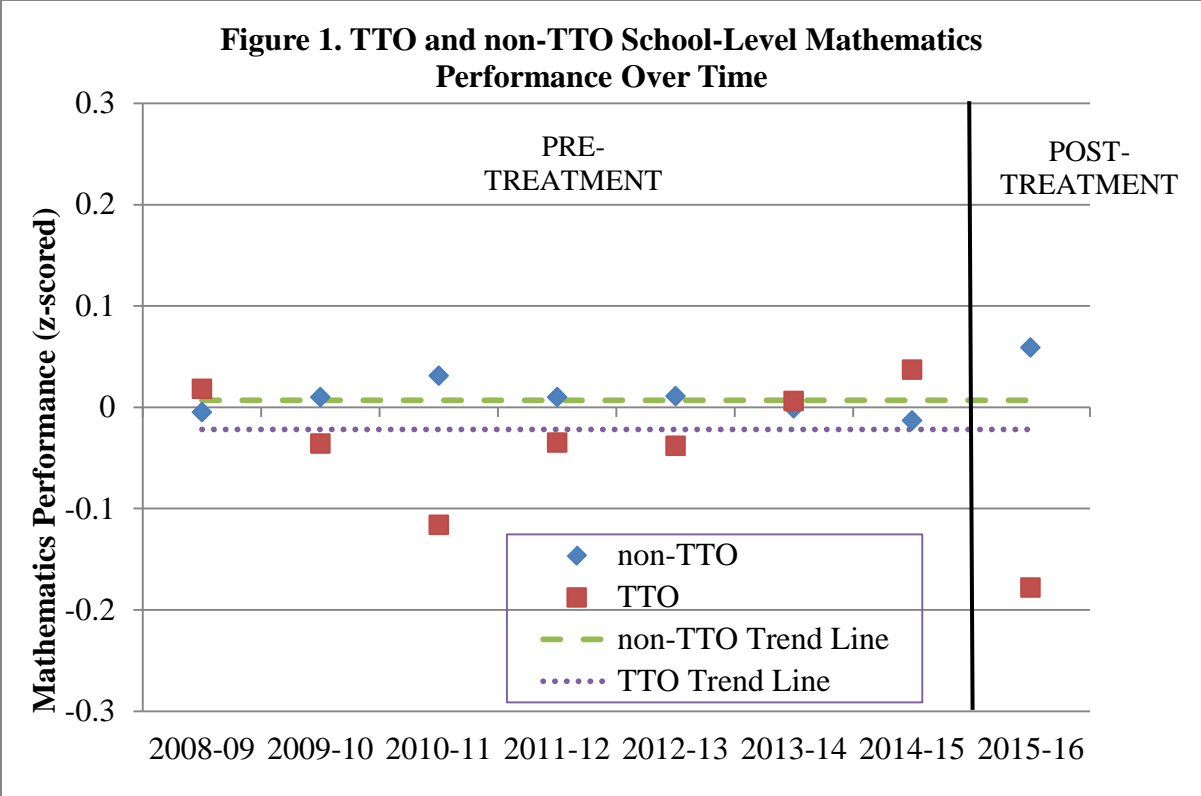**Figure 1. TTO and non-TTO School-Level Mathematics Performance Over Time**

Table 2 presents the actual results of the CITS analyses for both the overall sample and individual grades. Our focus here is the estimated TtO treatment effect in year 1 of implementation, represented by the cross-level interaction of year 1 (the cohort level year indicator) by TtO (the school-level indicator of TtO participation). As indicated in the far right column, across all grades combined, we find no significant effect of TtO in year one ($p>.10$). However, the TtO estimates vary somewhat across grade levels. With fifth grade, these results suggest a marginally significant negative impact of TtO participation on student mathematics performance ($ES = -0.371$ SDs; $p<.10$). We find no significant effects of TtO in sixth through eighth grades ($p>.10$).[8]

We should also note that compared to Elizabeth Public School students in the same cohort within the same school, regardless of whether they participated in TtO, females slightly

---

[8] Because TtO is a school-level intervention, statistical power—our ability to identify a statistically significant impact—is strongly determined by our school-level sample sizes. Given the relatively small number of both treatment and control schools employed in these models, we use the more liberal $p<.10$ level of statistical significance for the school-level estimates. We do the same with the HLM gain-score models discussed below.

13

underperform their male peers in mathematics, while language minority, special education and free-lunch status (in particular grades) are negatively associated with math achievement.[9] Asian students exhibit somewhat higher mathematics performance compared to white students, while black and Hispanic students have lower average test scores compared to their white peers. To capture changing demographics both within and between EPS schools over time, these models also include aggregated indicators of these student-level measures. We see that few are substantively or significantly associated with student mathematics performance, once we account for the student-level associations.

---

[9] Again, these estimates are not exclusively for TtO students. Rather, they are for all students *controlling* for TTO status, as well as the other cohort- and school-level controls.

Table 2. CITS (Difference-in-Difference) Estimates of TtO Impact on Mathematics Achievement in Elizabeth, New Jersey, Year One

	Fifth Grade (n=7,742)	Sixth Grade (n=8,443)	Seventh Grade (n=7,935)	Eighth Grade (n=7,614)	**OVERALL (n=31,734)**
Students (Level 1)					
Female	-0.058**	-0.043*	-0.066**	-0.096***	-0.067***
Limited English Prof.	-0.657***	-0.683***	-0.704***	-0.729***	-0.695***
Special Education	-0.616***	-0.783***	-0.827***	-0.728***	-0.759***
Free Lunch[1]	-0.049	-0.088*	-0.066	0.071*	-0.027
Reduced-Price Lunch	0.090	0.087	0.071	0.196***	0.122
Black[2]	-0.377***	-0.440***	-0.472***	-0.421***	-0.433***
Hispanic	-0.084*	-0.174**	-0.224***	-0.163***	-0.163***
Asian	0.293**	0.201*	0.056	0.134	0.168***
Fifth Grade[3]	--	--	--	--	-0.038*
Sixth Grade	--	--	--	--	-0.016
Eighth Grade	--	--	--	--	-0.005
Cohorts (Level 2)	(n=127)	(n=143)	(n=143)	(n=143)	(n=143)
% Female	-0.001	0.000	0.001	0.001	0.000
% Limited Eng. Prof.	-0.002**	-0.002**	-0.003***	-0.002***	-0.002***
% Special Education	-0.001	-0.001*	-0.001	0.000	0.000
% Free Lunch	-0.001	0.000	0.000	0.000	0.000
% Red.-Price Lunch	0.001	-0.001	-0.002*	-0.001	-0.001
% Black	-0.002	-0.001	-0.001	-0.003*	-0.002*
% Hispanic	-0.001	0.000	0.000	-0.001	0.000
% Asian	-0.001	0.000	0.002	-0.002	0.000
Year 1	0.008	-0.060	-0.043	-0.082	-0.043
TtO X Year 1	**-0.371~**	**-0.036**	**-0.047**	**-0.166**	**-0.121**
Schools (Level 3)	(n=18)	(n=18)	(n=18)	(n=18)	(n=18)
% Female	-0.003	-0.005**	-0.003*	-0.001	-0.003*
% Limited Eng. Prof.	-0.002	-0.002	-0.003	0.000	-0.001
% Special Education	-0.003*	-0.002**	-0.001	-0.001	-0.001*
% Free Lunch	-0.002	-0.004**	-0.002	0.000	-0.002~
% Red.-Price Lunch	0.001	-0.005~	-0.001	0.008~	-0.001
% Black	0.001	0.002*	-0.002	0.002	0.002
% Hispanic	0.001	0.003*	0.002~	0.002	0.002~
% Asian	0.002	0.006*	0.002	-0.006	0.001
TtO	-0.134	-0.081	-0.034	-0.060	-0.059
Intercept	0.020	0.024	0.021	0.027	0.017

~p<.10; *p<.05; **p<.01; ***p<.001. Outcome is z-scored (standardized) within grades and years.
[1] Free and reduced-price lunch compared to full-price lunch students.
[2] Racial/ethnic comparisons to white students.
[3] Grades compared to seventh grade.

<u>MAP Results</u>

Descriptive analyses. Figures 2-5 display MAP results by grade and TtO status. Fifth-grade TtO and non-TtO students began the school year with comparable math skills (see Figure 2). However, non-TtO fifth-grade students gained somewhat more skills during both the fall-to-winter and winter-to-spring testing periods. This divergence between TtO and non-TtO fifth graders during the first implementation year is reflected in the CITS results discussed above. Sixth-grade TtO students began the academic year at a slight advantage compared to their non-TtO peers, but gained skills as a slightly slower rate, resulting in no skills difference by the winter. Between winter and spring, TtO and non-TtO sixth-grade students had virtually identical learning rates. We found similar patters with both seventh (Figure 4) and eighth grade (Figure 5): slower growth rates during the initial implementation period, but comparable growth rates in the subsequent period.

Figures 6-9 divide students within each grade into thirds based on their initial (fall) MAP scores: students in the bottom third of the distribution for their grade are designated "low-achieving," those in the middle third "average-achieving," and students in the top third of the grade-specific distribution are labeled "high-achieving". Within each initial-achieving strata, TtO students gained fewer math skills than their non-TtO peers during the fall-to-spring period. The fact that the TtO/non-TtO growth differentials are smallest among initially higher-achieving students stems largely from the fact that mathematics growth on the MAP assessments is negatively associated with initial achievement status—higher-achieving students gained less during the academic year.[10] Indeed, among these Elizabeth students, we find a sizable negative association between fall MAP achievement and fall-to-spring mathematics gains ($r = -.481$; $p<.001$). In recognition of this phenomenon, we turn now to analyses that estimate growth rates among TtO and non-TtO students while accounting for their initial mathematics skills.

---

[10] For an analysis of this issue with MAP conducted by NWEA staff, see McCall, M.S., Hauser, C., Cronin, J., Kinsbury, & Hauser, R. (2006). *Achievement gaps: An examination of differences in student achievement and growth*. Lake Oswego, OR: Northwest Evaluation Association.

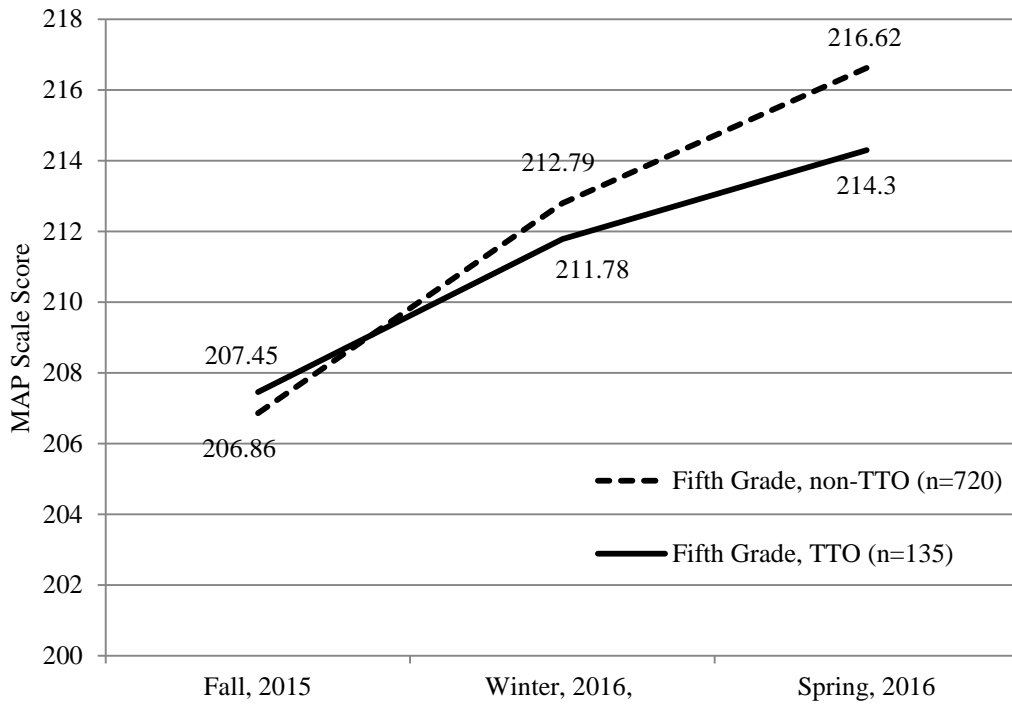Figure 2. TtO and non-TtO Fifth Grade MAP Math Growth
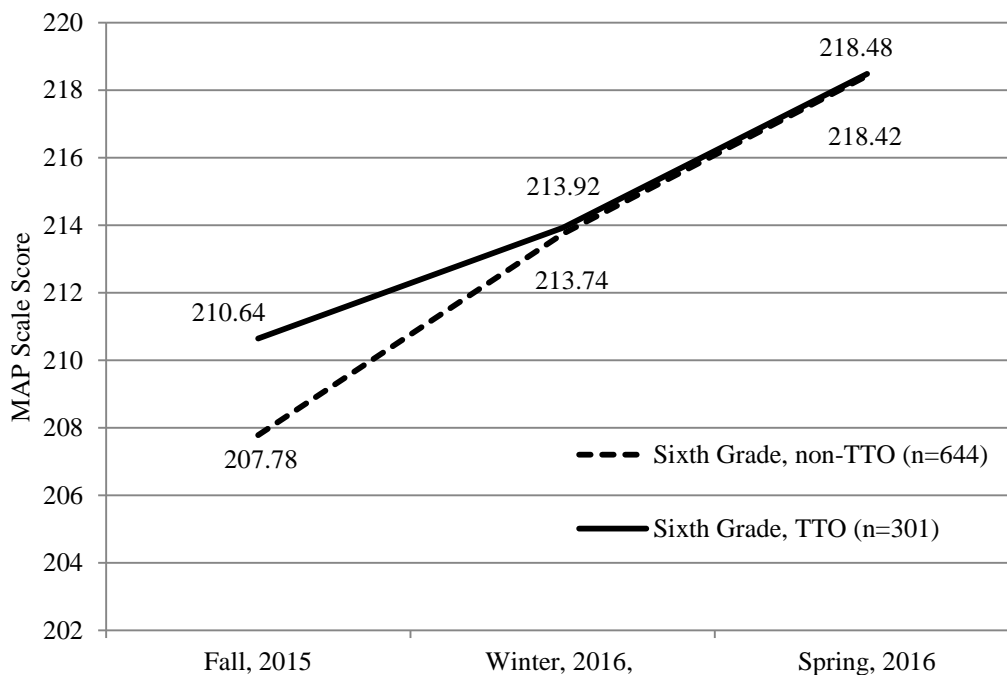

Figure 3. TtO and non-TtO Sixth Grade MAP Math Growth

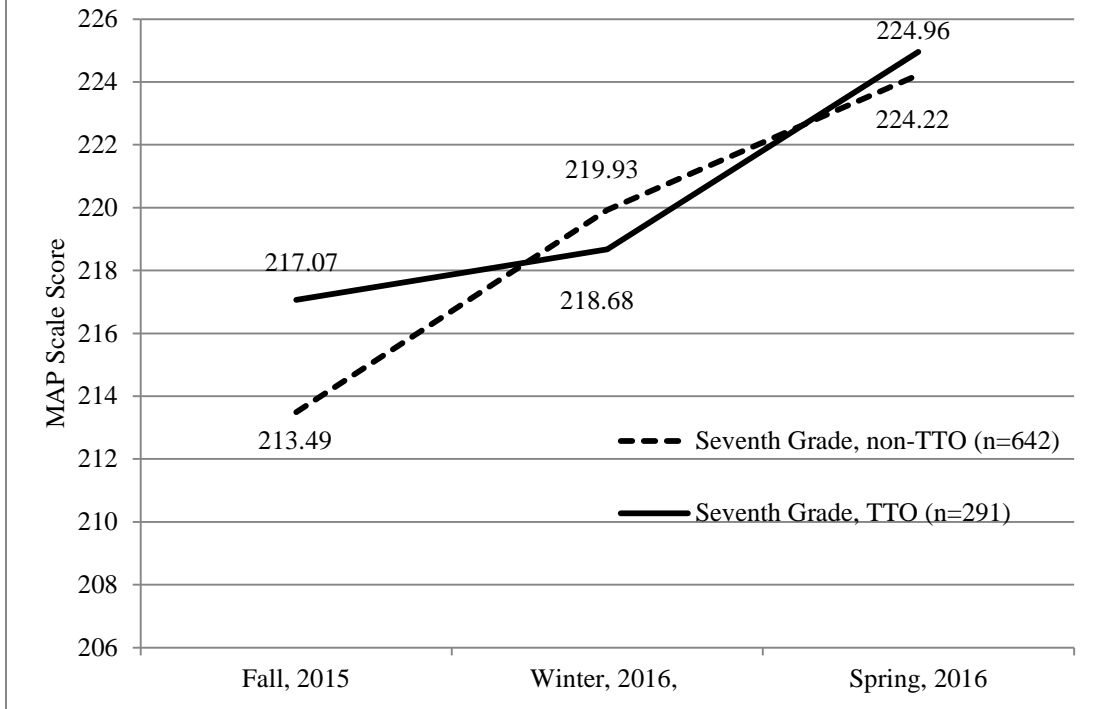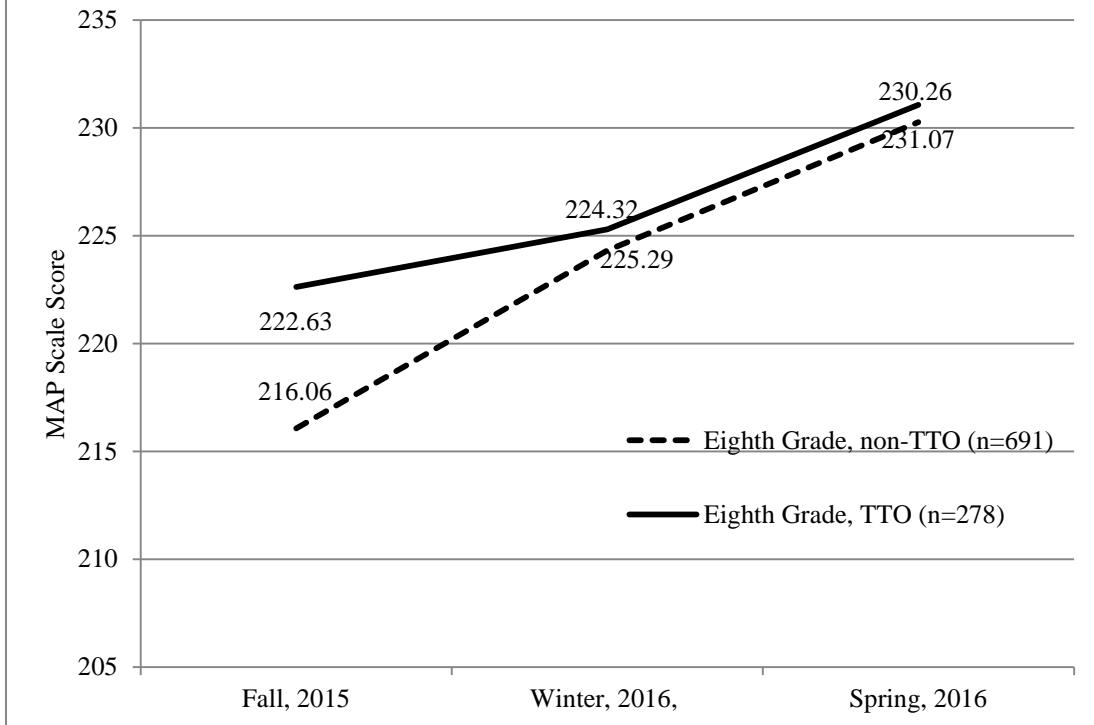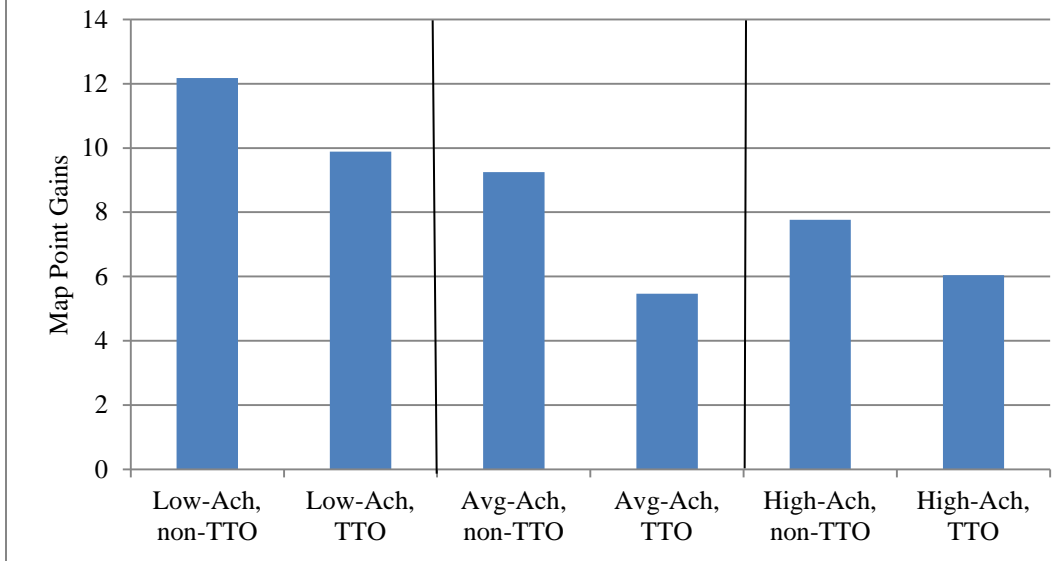Figure 4. TtO and non-TtO Seventh Grade MAP Math Growth

Figure 5. TtO and non-TtO Eighth Grade MAP Math Growth

Figure 6. Fifth Grade MAP Math Growth by Fall MAP Achievement

Category	Map Point Gains
Low-Ach, non-TTO	12.1
Low-Ach, TTO	9.9
Avg-Ach, non-TTO	9.2
Avg-Ach, TTO	5.4
High-Ach, non-TTO	7.8
High-Ach, TTO	6.0

Figure 7. Sixth Grade MAP Math Growth by Fall MAP Achievement

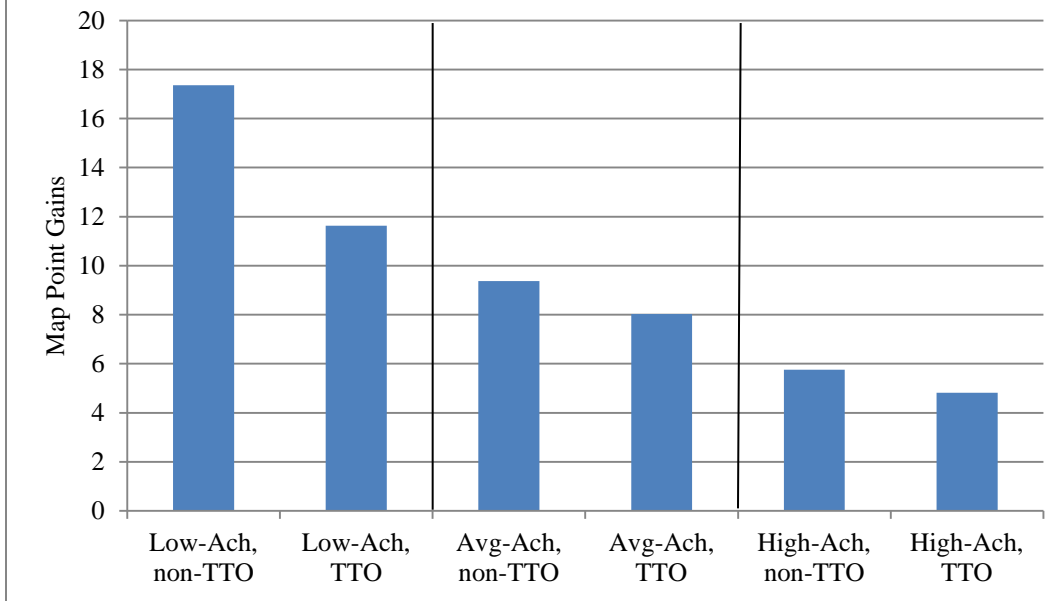| Category	Map Point Gains
Low-Ach, non-TTO	17.3
Low-Ach, TTO	11.6
Avg-Ach, non-TTO	9.3
Avg-Ach, TTO	8.0
High-Ach, non-TTO	5.7
High-Ach, TTO	4.8

Figure 8. Seventh Grade MAP Math Growth by Fall MAP Achievement
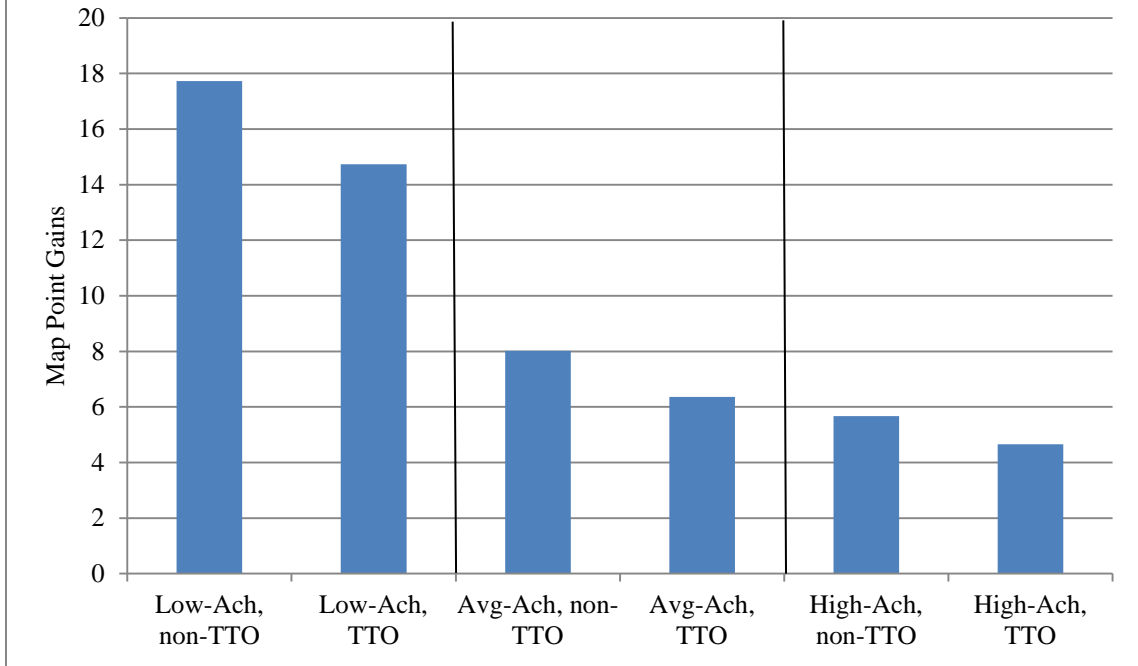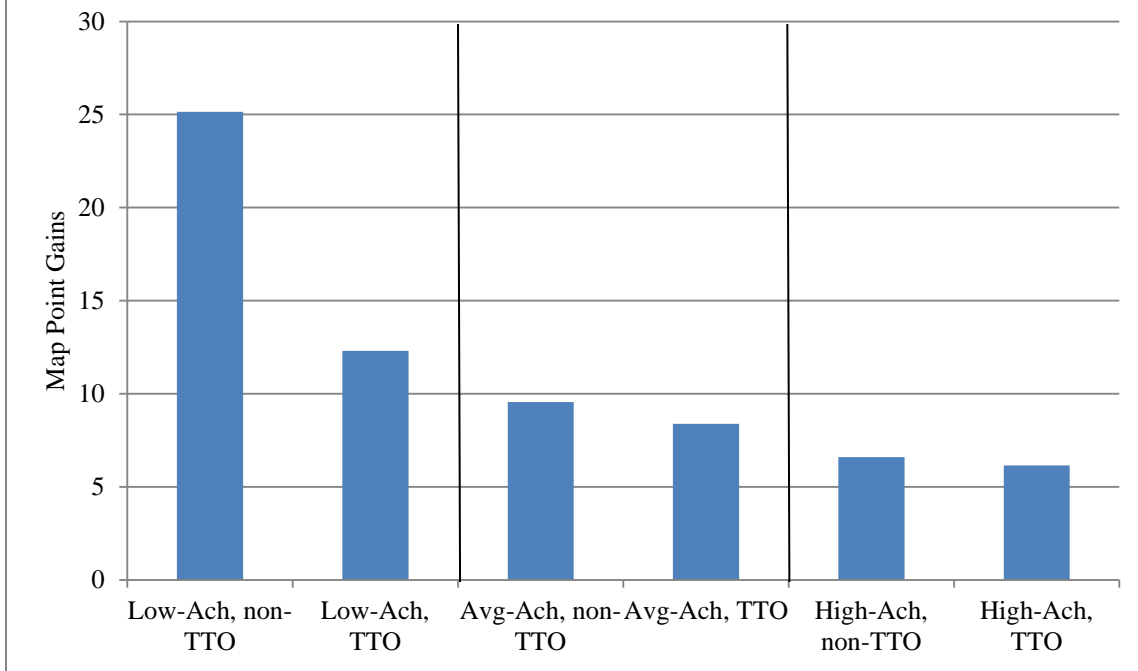

Figure 9. Eighth Grade MAP Math Growth by Fall MAP Achievement

HLM gain-score analyses. Model 1 in Table 3 presents the estimated relationship between TtO school status and student mathematics learning during the 2015-16 academic year, adjusted only for initial (fall) mathematics ability. We find that students in TtO schools gained somewhat fewer mathematics skills compared to their peers attending non-TtO schools (*ES* = -0.280; *p*<.05). The moderately large negative estimate associated with fall math achievement suggests the importance of controlling for the relationship between initial skill status and subsequent skill growth. We find that initially higher-achieving students exhibited substantially weaker gains during the academic year (*p*<.001), regardless of whether they attended a TtO school. As noted above, this "fan-close" phenomenon has been reported elsewhere with student-level MAP data.

Model 2 adjusts the TtO estimate for student-level social and academic background characteristics. Given that the TtO and non-TtO schools in this sample enroll students with quite comparable backgrounds (see Table 1), the TtO estimate should change very little once we take student backgrounds into account. In other words, although we would expect associations between student characteristics and mathematics learning, those characteristics are equally distributed across TtO and non-TtO schools. Indeed, the TtO estimate from Model 2 changed only slightly from the previous model (*ES* = -0.255; *p*<.05). The student-level estimates closely resemble those presented in Table 2, with females marginally underperforming their male peers, and larger disparities in mathematics growth associated with language minority and special education status. Moreover, black students also exhibited lower average test score gains compared to their white peers, while Asian students gained more mathematics skills, all else equal. Grade level was unrelated to mathematics learning. We also constructed models with cross-level interactions to explore whether the TtO estimate varied across grades. These models indicated that it did not.

Model 3 introduces the school-level indicators, which are simply aggregates of the student-level measures. We again find the TtO estimate to be virtually unchanged from the previous model. The student-level coefficients are similarly robust across models. Given that the school sample was created to have as little variability across schools, the lack of significant school-level estimates is both welcome and unsurprising.

Model 4 incorporates a cross-level interaction between student's fall mathematics achievement and TtO participation. The TtO main effect is now the estimated TtO impact for

students with average initial math achievement (*ES* = -0.283; *p*<.05). The significant interaction term suggests that the TtO effect varies by students' initial math skills. Specifically, among higher achievement students (e.g., those with initial achievement one standard deviation above the mean), the negative TtO effect is largely eliminated (i.e., the -0.283 SD TtO main effect is substantially reduced by the 0.214 SD interaction estimate). Among initially lower achieving students, however, the negative effect of TtO is stronger (e.g., [-0.188 + (-0.215)]). This is precisely the relationship displayed in Figures 6-9: wide disparities in mathematics growth between low-achieving TtO and non-TtO students, but very small differences between high-achieving students, regardless of their TtO status.[11]

Winter to Spring gains. The gain-score models above include the challenging startup period in the Fall of the 2015-16 year. These initial challenges are clear in Figures 2-5. We constructed models that used winter-to-spring growth on the MAP math assessments as the outcome, controlling for initial MAP (winter) math scores. These models clearly support Figures 1-4, with parallel growth rates among TtO and non-TtO children (*ES*=0.031; *p*>.8).

---

[11] We also tested whether TtO had differential effects for LEP and SPED students. We found that it did not, with neither slope varying as a function of TtO status (*p*>.10).

Table 3. TtO and Student Mathematics MAP Gains in Elizabeth, New Jersey, Year One
(n=3,780 students in 5 TtO and 13 non-TtO schools)

	Model 1	Model 2	Model 3	Model 4
Students (Level 1)				
Female		-0.071*	-0.072*	-0.056*
Limited English Prof.		-0.291***	-0.292***	-0.203***
Special Education		-0.210***	-0.207***	-0.132**
Free Lunch[1]		-0.010	-0.010	0.017
Reduced-Price Lunch		-0.032	-0.032	-0.020
Black[2]		-0.139*	-0.136*	-0.119
Hispanic		0.007	0.002	0.017
Asian		0.260*	0.258*	0.327**
Fifth Grade[3]		-0.006	-0.004	0.018
Sixth Grade		0.002	0.002	-0.008
Eighth Grade		-0.015	-0.016	-0.042
Fall Math Ach.	-0.448***	-0.496***	-0.496***	-0.520***
Fall Math Ach.*TtO				0.214*
Schools (Level 2)				
TtO	**-0.280***	**-0.255***	**-0.275***	**-0.283***
% Female			-0.001	0.001
% Limited Eng. Prof.			0.004	0.004*
% Special Education			0.000	0.001
% Free Lunch			0.000	0.001
% Red.-Price Lunch			0.001	-0.001
% Black			0.001	-0.001
% Hispanic			-0.002	-0.001
% Asian			0.000	-0.005~
Intercept	0.068	0.188*	0.193*	0.136

~$p<.10$; *$p<.05$; **$p<.01$; ***$p<.001$. Outcome is z-scored (standardized) within grades.
[1] Free and reduced-price lunch compared to full-price lunch students.
[2] Racial/ethnic comparisons to white students.
[3] Grades compared to seventh grade.

Robustness Check: ELA Results

Arguably, the TtO model should not influence—either positively or negatively—student ELA performance. As an additional robustness check, we conducted analyses identical to those described above, but with state ELA test scores as outcomes for the CITS models, and MAP ELA gain-scores as outcomes with the two-level HLM growth models (adjusted for initial ELA achievement). We did so because potential effects of TtO on ELA scores that were similar in magnitude to those reported above for mathematics would obviously raise concerns. Namely, that the TtO schools were, in the initial TtO implementation year, simply less effective for

23

whatever (unmeasured) reasons. However, a finding of negative effects of TtO on student math performance, but not on ELA test scores, would support the finding of negative program effects in year one. This is indeed what we found. The results of these additional analyses suggest no effects of TtO on student ELA performance with the CITS models (ES=-0.04; p>.5) or with the MAP gain-score analyses (ES=0.06; p>.4).

PROCESS EVALUATION

Data and Methods

CPRE conducted qualitative field research in Elizabeth, NJ during February, 2016. We visited all five *Teach to One: Math* schools, observed classes, and interviewed a total of 44 participants, including five school principals, two vice-principals, five Math Directors, 16 teachers (Math Advisors), five Teacher Residents (TRs) for special education students, five Teacher Residents (TRs) for bilingual students, and six New Classrooms staff members (two Site Operations Managers (SOMs), two Operations and Technology Associates (OTAs), and two coaches; see Table 4 below). Interviews were semi-structured and lasted approximately 30-45 minutes each.

Table 4. Qualitative Field-Work Participants, 2016 (*n*=44)

School	Principals	Math Director	Teacher (MA)	TR-SPED	TR-Bilingual
3	1	1	4	1	2
18	1	1	5	1	0
21	1	1	2	1	1
23	1	1	3	1	2
28	1 (+2 VPs)	1	2	1	0
Total	*7*	*5*	*16*	*5*	*5*

We analyzed these qualitative data using the *Atlas.ti* software. CPRE researchers coded each interview separately and analyzed response patterns across the interviews. The following pages synthesize the themes revealed through this fieldwork and subsequent data analysis. The sections below address multiple topics, including staff perceptions of: 1) how TtO is working for them; 2) the quality of TtO's curricular and instructional content; 3) how the model is working for students, and; 4) the technical and instructional support provided by New Classrooms.

Process Results

I. How is the TtO model working for teachers?

This section explores feedback CPRE researchers received from teachers implementing the TtO model, including the extent to which TtO teacher assignments were aligned with teacher background, how teachers felt about their allotted preparation time, to what extent teachers

collaborated or employed their common planning time, how teachers felt about the new joint teacher evaluation system, and how teachers were reacting to their new roles in the classroom.

Tailoring Teacher Assignments to Teacher Background

Numerous teachers mentioned feeling that their backgrounds were not always conducive to teaching their assigned lesson; this was particularly the case with Live Investigations but pertained to other modalities as well. For example, one teacher stated that she struggled to teach lower-level content (such as basic division) in a way that was accessible to students, as this teacher had been teaching higher-level mathematics for many years. Another teacher echoed this sentiment, stating that they "struggled with how to make the concepts interesting (for lower grades) because I don't have experience teaching the content." Likewise, teachers who previously taught lower grades also felt that they sometimes struggled with the higher-level content; some teachers reported feeling increased stress-levels associated with being assigned to content beyond their comfort level, coupled with limited preparation time.

Overall, some teachers felt that the program "has not taken advantage of [their] pedagogical background." One teacher asked, "Why not use our capacities? I think we should be placed where we feel comfortable with the content." In general, some teachers expressed that it had been difficult to add the "higher level of creativity" to their lessons (compared to past years), due to lack of preparation time and the fact that they were teaching so many grade levels in a single day. As one teacher said, "I used to teach just one grade, and now I have up to three others grades whose skills I need to brush up on." As a "work-around," one coach mentioned that they encouraged teachers to "switch" roles if their backgrounds were better suited to certain topics or grade-level material. This coach argued that "teachers need to take ownership" over their assignments and "feel free to switch things up"; other teachers mentioned that this did happen informally in their classrooms but only on an ad-hoc basis.

Teacher Preparation Time

The vast majority of teachers felt that their planning time was reduced under TtO compared to previous years. Although many appreciated the reduced prep-time required, a majority felt that this abbreviated amount of planning time was inadequate. Many teachers stated that they had too little time to prepare for their courses (particularly Live Instruction), as they

26

received their assignments too late the night before (between 5:30 and 8:00 p.m.). Teachers who had first period "prep" found that the time allotted for preparation was much more manageable. Even teachers teaching Small Group Collaborations stated that they would like to be more prepared to work with their students (i.e. receive all the assignments ahead of time), although they realized that the TtO format precluded this. A couple of teachers requested a "teachers manual" covering all topics spanning 5th through 8th grade that would provide them with a quick and easy way to brush up on particular skills. Teachers stated that they were always "playing catch-up" and that the limited time they had to prepare "limits their ability to pull from different resources" and create a more engaging lesson. Teacher Residents (TRs) without content knowledge felt particularly under-prepared to help multiple students in the same modality working on many skills simultaneously. Teachers also wished for more time between activities/rounds/cohorts as well, in order to have time to "process" and "recharge." Of course, this desire is shared by those teaching traditional courses back-to-back during the school day.

Some more-experienced teachers felt that they had an "artillery of tools" that made it easier to prepare lessons on the spot, but that a "novice may be more receptive to the program generally but also doesn't have the strategies or experience needed to bring lessons to life, even in a virtual setting." Multiple first-year teachers did indeed feel that this abbreviated preparation time was particularly hard for them, as they needed more time to plan and explore practices that worked for them.

Teacher Collaboration & Common Planning Time

Teachers reported mixed impressions of collaboration opportunities and common planning time. One teacher expressed the enjoyment she found working with other teachers within the TtO format, noting also that her school had a history of collaboration that included "letting each other know which students were struggling the day before" so that those students could get more attention throughout the various modalities. Other teachers claimed that they shared Live Investigation lesson plans and often talked about best practices.

Teachers who did not have common planning or preparation time expressed that they collaborated much less. Importantly from the standpoint of establishing fidelity of implementation, the amount of common planning time varied drastically across schools, ranging from none, to two or three days per week, to every day. In some schools, common planning time

for special education teachers was separate, meaning that these teachers did not have the opportunity to collaborate with non-SPED teachers. In another school, although common planning time was officially scheduled, some teachers taught other classes during this period.

Teacher Evaluations

Teachers held a variety of views regarding the new accountability approach within the TtO framework, in which teachers were collectively (rather than individually) responsible for all students in the TtO program. Roughly one-third of teachers viewed this new communal approach in a positive light. For example, one teacher stated that they preferred to "sink or swim" together; another felt that the TtO teaching team was so strong that "good collaboration" and improvements in student achievement were within reach. Others claimed that they "trusted" the teachers they worked with, and that if "everyone does their job, we will be fine; in this building we work together and support each other."

However, roughly one-third of teachers interpreted the new approach to teacher accountability in a negative light, with one expressing that she felt "uncomfortable owning the results of all 200 students, including some I may or may not see every day." Another lamented that she felt "guilty" if she did not manage to "get students to understand the material," and that she worried about how that will affect other teachers.[12] This same teacher, however, questioned the fairness of evaluating teacher quality collectively when not all teachers were certified in math, and assignments to Live Instruction were not necessarily based on teacher strengths or backgrounds. Others stated that they did not have enough "faith" in their colleagues, and since the evaluations were based on PARCC results, it was unclear how much student performance would reflect the hard work teachers put into supporting the TtO program (which they argued had limited focus on PARCC content).

The final third of teachers were neutral about the new joint evaluation system, viewing it as "part of signing onto TtO," for better or for worse, but that it was worth it to be part of a program that "meets students where they are." Interestingly, teachers also asserted that there was more performance pressure associated with the Task modality, as student performance is associated with that specific Task teacher.

---

[12] To supporters of teacher accountability systems, this is, of course, a positive outcome.

Teacher Roles and Accountability

Many teachers lamented that the TtO model decreased their autonomy. Rather than working independently, they now functioned as part of a team with other teachers. Some felt that this caused them to lose their "identity" or "character." For example, as one teacher reported, "Teachers are used to being captains of their own ship, but when you're doing TtO, you're part of a team, and you have to be able to leave your ego at the door." This feeling of dislocation was particularly acute for TRs, some of whom felt "like a babysitter," that they were "totally removed from teaching," or that being a TR is "boring." Several teachers expressed disappointment with losing their own physical space, while others felt that they still retained their own space that they could control.

Numerous teachers claimed that the TtO model made it more difficult to build relationships with students, largely because they did not work with the same groups of students for sustained periods every day. However, some teachers felt that they were able to build relationships with their Math Advisory students. "I see them every day. If they get in trouble, I feel like they're mine." Some teachers reported doing their "own individualization of the homework" for their MAs, or inviting their MAs to come at lunch for extra help. Some teachers also shared that MAs were a useful time for answering students' questions. However, other teachers said that they did not have one-on-one conversations with the students in their MAs, did not check in with their MA students about their progress, or did not have enough information about their MA students' learning progressions to be able to help them effectively.

Many teachers reported frustration with their inability to follow up with students or deliver feedback. Although some expressed satisfaction with being able to see student mastery via exit slips, many teachers wished they could have more continuity with students and over time be able to follow up with them into the next period or the next day, especially for those students who "maybe… have 80% of a concept, and they are missing one component." Similarly, some teachers wished that they had more time to go over homework with students instead of just passing it back or grading for completion. "Unfortunately, as I go over homework and I see mistakes, I can't go over it. The program doesn't give me time… I feel horrible that I can't go over homework with them."

Teachers were divided as to who was ultimately responsible for student learning. All respondents agreed that students bore a certain amount of responsibility, but some teachers

believed that students alone were ultimately accountable for their learning, while others stated that students and teachers shared accountability. A smaller number felt that accountability was shared by all stakeholders, including students, teachers, and New Classrooms. Teachers generally believed that the TtO model provided students "all the resources" they needed to be successful, should they choose to take advantage of them. "They [students] have the tools and the opportunity to ask us questions and for help." Some teachers also reported that "students have become more independent through the program," which they felt would be useful later in college. Although many teachers felt that they should also be held responsible for student results, a minority expressed feeling "uncomfortable with owning the results of all 200 students, including some who you may or may not see every day."

II. How do schools feel about TtO's instructional content?

This section details the feedback provided to CPRE regarding schools' impressions of TtO's instructional content, including views regarding each modality (Task, collaborative modalities: Peer to Peer (P2P) and Small Group Collaborations (SGC), Live Investigations (LINs), and Independent Learning Zones (ILZs), as well as interpretations of TtO's pacing and timing, homework, instructional content, algorithm, portal, and assessments. This section concludes with an examination of the extent to which schools felt that TtO prepares students for state standardized tests.

Task Modality

In general, teachers valued the theoretical concepts behind and practical manifestation of TtO's Task modalities. They enjoyed seeing a group of students proceed through a set of skills for several days in a row. They also appreciated that the Tasks involved "real-world applications" and that they have a longer planning period before a Task. However, some teachers noted that not all Tasks met the "real world applications criteria" and that some had to be adapted or modified "to make it attainable for all groups." Other teachers wished that the Tasks were "less wordy" and "less language intensive" in order to be more accessible and digestible to students. One teacher also noted that material from days 1 and 2 that has not yet been mastered will often get pushed to day 3 of Task, as the exam on day 3 tests day 1 and 2 skills; for this reason, teachers felt that they sometimes ran out of time to cover all the skills before the Playlist

demo and that there was frequently insufficient time for students to present their work. From the limited number of classroom observations conducted by CPRE staff (four observations of a Task lesson), researchers found that most students were working independently on their Task most of the period (after the teacher modeled an example), while occasionally checking in with their peers. In three of the four classes observed, the Task instructional space was relatively noisy with a handful of students off-task at various points in the lesson.

Peer-to-Peer (P2P) & Small Group Collaboration (SGC)

While multiple teachers appreciated the idea of P2P and SGC as a way for students to work together to solve problems and to assume leadership roles, most teachers felt that these collaborative modalities were not as successful as they could be. Teachers stated that the worksheets were too long and that students could benefit from shorter worksheets with fewer, but harder problems. They also suggested adding examples to the worksheets, as a number of students come to these collaborative sessions without having mastered pre-requisite skills, or were seeing the skill for the first time. In addition, the fact that groups of students work on different packets may make it harder to prepare for and facilitate the sessions.

Teachers also stated that in some instances pairing "high" and "low" ability students had led to productive sessions where students taught and learned from each other. But in other circumstances, such pairings permitted less-motivated students to simply evade the assigned work. Teachers also described struggles to make students work together when some preferred working independently, while others, such as English-Language Learners and Special Education students, needed more time. They also stated that certain students ended up using the time to simply chat with each other. In general, teachers asserted that the quality of the session depended in part on how well the teacher was able to manage the small groups and their engagement levels.

Overall, teachers felt that these modalities could be improved by introducing a greater level of accountability beyond the behavior score, which some teachers saw as an empty threat. Teachers suggested that students should have to "turn in something" or show the teachers their results online. In some schools, teachers assigned roles for each student within the SGC group, which seemed to add structure and accountability to the modality, but this practice was only employed in two schools. Finally, teachers mentioned that the P2P and SGC modalities were not

clearly structured for peers or small groups; for example, one teacher stated that when creating such work, "you must create inter-dependencies across students so that they have to rely on their peers … right now, students can do the assignment on their own – and then they come together and talk about it, but the structure does not necessarily facilitate group discussion."

Another teacher suggested Virtual Instruction and P2P be combined in such a way that students first watch a video together and then discuss the content as a lead-in to the assignment. CPRE Researchers conducted a limited number of observations of P2P and SGC modalities (three in total). We observed wide variation in the extent to which students worked collaboratively, even within individual classrooms. In each observation, roughly half of the class (or half of the groups) were working independently while the other half was working collaboratively. Teachers had to encourage students to work with their peers saying, "This is supposed to be a collaborative assignment; you must work together." It was unclear why certain groups were working predominately independently; in some cases it may have been due to language difficulties or the pairing of students whose skills were not complementary.

Live Investigations (LIN)

Teacher feedback on Live Investigations was overwhelmingly positive. Teachers enjoyed interacting with and providing more personalized live instruction to student groups that were much smaller than those found in a regular classroom. Teachers further asserted that this modality was generally a student's introduction to a skill and often "provided the foundation for the skill." Teachers also appreciated the freedom to base their lesson on the TtO template or to create a lesson on their own (though a number of teachers felt that they did not have sufficient preparation time for this; see "Teacher Preparation Time" section above).

Despite the strong support for the modality, teachers did offer suggestions for how to improve the LIN modality. Multiple teachers claimed that some LINs contained too many skills or required pre-requisite skills that not all students possessed, causing the LINs to feel rushed if a teacher needed to scaffold the material afresh; others stated that not all students benefited equally from LINs and that "outlier groups such as students performing very high or very low" are getting "left out of the LINs," and that it did *not* seem to be due to the fact that there were too few students on a skill. Finally, some teachers reported that it was difficult to personally follow-up with students later that day or the next day in order to provide the final instructional pieces

that might solidify a skill for them if the LIN was not long enough. From the limited number of classroom observations conducted by CPRE staff, researchers reported that students seemed engaged in the Live Investigations modalities and that teachers had time to continuously check student written work (by circulating in the space) and to work with students individually.

Independent Learning Zones (ILZ)

Teachers were relatively divided as to the effectiveness of the Independent Learning Zone modalities. Roughly one-third of teachers found ILZs to be effective, pointing to the quality and diversity of the online material, and the fact that students can hear the information multiple times and rewind as necessary. These supportive teachers also claimed that this modality increased student ownership of their own learning. The remaining two-thirds of teacher offered mixed views of the ILZs' effectiveness, which they felt varied across student type. More "motivated" and "disciplined" students, these teachers argued, got much more out of these modalities, while other students simply "played games" and treated ILZs as "free periods." Some of these teachers reported that it was difficult to ensure that all students were engaged and working or fully understanding the material, and that some students feigned understanding out of either embarrassment or a desire to be left alone. For example, one as teacher stated, "The independent stations require students to be responsible enough to focus on what they are doing… but middle-schoolers do not always feel comfortable to say that they don't know what they are doing… some are not independent enough to be in this area or are not humble enough to ask for help."

Further, some teachers expressed that the level of support provided to students through the online platforms was inadequate. As one teacher claimed, "If students have a follow-up question, the computer won't answer it for them . . . even though the computer can give them hints and support, they don't think the hints and videos are helpful." Particularly for special education students, teachers worry that it may be difficult for them to "process what they read or see" and that they "need a live person to break it down for them." Across the board, all teachers stated that the ILZ modality could be improved by increasing student accountability, by either requiring all students to turn in their notes or their independent practice, or to share the results of their work online. Further, in several schools, teachers felt that there were not enough staff members to support the needs of students in the ILZs, particularly special education students.

Researchers observed only four lessons utilizing the Independent Learning Zone. Overall, researchers noted that teachers circulated around the room and focused their time largely on students who raised their hands or students who needed additional support, such as special education students or English Language Learners. In one lesson we observed, the teacher also used the iPad to determine which students might need assistance (students who were attempting a task for a third time) and pulled these students aside to work with them. Researchers reported that while they did witness off-task behavior, the majority of students were relatively quiet and working on their assignments.

Timing and Pace

Teachers had mixed views regarding the timing and pace of instruction. Some teachers noted that the shorter periods kept students from "getting bored" or "getting off task" and appreciated the fact that students could "move at their own pace" with "multiple opportunities" to master a skill. One coach noted that teachers needed to realize that their lesson was often an "introduction" and that even though 35 minutes (or less) per session is short, the students would have multiple opportunities to master the material.

However, other teachers felt that the timing, particularly for the Live Investigations, was too brief; teachers stated that even though they may be providing only an introduction, there was too little time to review past skills and build off of them. Though all students in a LIN were supposed to have necessary background knowledge, teachers stated that this was not always the case and that even for students who had "passed" the material before, all students benefited from "scaffolding" these new skills (reviewing, and in some cases re-learning, foundational skills and linking these to new content).

Even in Virtual Instruction, teachers felt that they often did not have time to reach all the students needing them. Teachers claimed that the time students needed to transition, log onto their portals, and complete their exit slips, left only about 25 minutes for instruction per session, which teachers felt was too short. One teacher also asserted that good classroom practice should involve discourse around both higher and lower order questions, but that under TtO there was pressure to "deliver the skill and content," and that it was "harder to find time to ask students to apply their skills or to think critically." This teacher admitted that she might not have fully

"wrapped her head around the TtO model yet" and that with time maybe she could adjust to the model.

Homework

There appeared to be great variation across schools in how homework was graded, with some schools grading for completion, some grading for accuracy, and others grading for a combination of the two. Across schools that graded for accuracy, the number of questions graded varied as well. Teachers like being able to grade for accuracy, as it shed a spotlight on student understanding, but they saw the process as quite time-consuming, as homework assignments differ with each student. Teachers also stated that they would like to go over homework with students the next day, but that their MA sessions did not provide time for this. One teacher instructed students to look through their homework on their own for feedback, although students rarely did this. One OTA stated that teachers often complained that there was little time for teachers to provide feedback. In addition, teachers were forced to translate homework by hand for English Language Learners, as the homework is a print-out and thus cannot be sent to Google translate.

Instructional Content

The vast majority of teachers and principals viewed the instructional value of TtO as resting with the program's promise to "meet students at their level." Teachers appreciated its ability to accommodate various student learning styles, provide instant feedback through exit slips, ensure good foundational knowledge, and challenge students at all ability levels. Teachers further felt that students could learn at their own pace, were more aware of their own progress, could learn independently, but still had opportunities to ask for assistance when needed. Teachers also appreciated the differentiated nature of the curriculum, given how difficult and time-consuming it is to differentiate instruction.

In terms of instructional shortcomings, teachers stated that it was often difficult to assess the quality of the online content, as teachers did not directly interact with this content unless they were supporting an individual student. Multiple teachers claimed that there were too few writing opportunities for students to explain their thinking and understanding. Teachers expressed a desire for more open-ended questions and opportunities for students to structure their responses.

35

Along these lines, numerous teachers mentioned the desire to assess a student's understanding through more than multiple-choice answers. One teacher stated that it was not until a student was on her fourth try that the teacher realized that the student was forgetting to simplify her fractions at the end, causing her to fail exit slips daily.

Finally, principals noted that it was challenging to assess the instructional value and performance of teachers under TtO. As one principal noted, "The difficulty is going in and observing them in this atmosphere and figuring out how this works with the Danielson framework; I am learning with the teachers; evaluating instruction isn't so black and white under TtO."

Algorithm Validity

All teachers interviewed stated that they could not truly assess the validity of modality and skill assignments, as the assignment process was not in their purview. One teacher stated that the assignment process felt like the "Wizard of Oz." A number of teachers stated that they "trusted" the algorithm and its ability to assign students to appropriate material, while others were somewhat unsure about the validity of the assignment mechanism and stated that at times it felt "sporadic," "wonky," or "curious." A few teachers mentioned that they had observed students struggling with skills above their grade level, and wondered how those students "got there." Other teachers stated that the skill assignments "jumped around a lot" between levels and seemed to re-assign certain students to skills that the teachers felt they had mastered. Teachers wished that they had access to a more developed concept and skills trajectory maps to know if the inconsistencies they were observing were due to different assumptions about student learning trajectories. Finally, some teachers expressed concern that grade-level skills tested on the PARCC were not prioritized in the TtO model, even for students on grade-level. Relatedly, numerous teachers requested additional algebra content.

Portal/Data

Across our sample, most teachers reported that they found the portal very useful because of the immediacy and accessibility of the data. As described in greater detail in the *TtO Testing* section (see below), TtO allows for data-driven instruction and very much informs how teachers plan their lessons. Although it is unclear whether parents and students were also utilizing their

portals, some teachers commented that they were helpful because parents could see for themselves how their children were progressing. Although individuals responded positively to the portal, they made a few suggestions to improve the experience. For instance, several teachers claimed that there were differences in teacher access across modalities. Unlike Live Investigations that have a teacher interface with notes and instructions, for Peer-to-Peer there is no teacher version, so teachers must take the time to work through the problems and troubleshoot. Furthermore, a few teachers wished for greater access to the data in order to generate reports for groups of students as opposed to looking at reports by each individual student, which they saw as time-consuming.

TtO Testing

Overall, teachers, principals and New Classrooms staff agreed there was a need for greater clarity and precision of TtO's grading system. Many respondents were curious about the grading algorithm and whether it was justified. Since TtO scores on a benchmark system rather than on a traditional 100-point scale, several interview participants (both teachers and principals) commented that the grades were misleading. They believed that the program was designed to make students appear more successful than they were and asserted that the level of work acceptable in the TtO model would not pass in a traditional classroom setting. Several attributed the grade inflation to the work ethic and work contribution components—daily checks on students' preparation and participation. They thought that the simplicity of requirements, such as bringing their supplies and raising their hand a few times, excessively boosted grades.

Furthermore, some individuals reflected on how TtO's grading system created a false reading of students' abilities because it is not indicative of student performance in terms of district- and state-wide norms. For example, a student may have received a 60% last year, and now may receive an 80%, but has been working on below-grade-level content. As one teacher put it, "The program is designed to set students up for success. You can't fail here. If you bring your notebook and pencil, then you're going to pass."

Although interviewees generally shared the same perspective on the grading system and what specifically needed improvement, their views regarding the use of exit slips in planning, how students reacted to these assessments, and the quality of the questions were more mixed. Overall, most teachers commented that exit slips were helpful for planning lessons. Being able to

see students' day-to-day progress allowed teachers to target students who had repeated the skill. A few teachers raised the issue that although the exit slips provided them with an opportunity for daily-generated data driven instruction, the very nature of students shifting modalities did not allow teachers to follow-up with the same students. However, one teacher explained that the data were still useful, "Whenever I finish a lesson, I look at the exit slip to see how the students did. I can't use the data to teach those same kids, since they won't come to me again for that content, but I can use it to change the lesson for when I teach it next time."

Respondents had mixed reactions whether the exit slips encouraged and motivated students. Some reported that students were competitive and compared exit slip results, while others stated that students knew that they were not graded so did not make an effort. According to teachers, student reactions appeared to vary by student; "There are some (students) that take it seriously and focus on what they're doing. There are others that get distracted. Some want to do better and persevere and stay afterwards, while others are just done after the exit slips." Moreover, teachers expressed concerns with the accuracy of measuring student learning. For one, they stated that multiple-choice questions do not always test understanding because students can guess and get lucky. They also suggested that some students used a calculator or accessed the browser for answers. Therefore, a lack of requirements for showing work undermined teachers' ability to determine whether students truly grasped the content.

TtO and Student Standardized Assessments

Conjecture on how the TtO model would affect student performance on standardized assessments included both hope and trepidation. Several interviewees believed that TtO would positively influence scores on the MAP, which assesses students at their respective levels, and would thus capture student mathematics learning. Other individuals commented that even if students were not at their appropriate grade level, they were at least learning the basic content required to advance; "This model better prepares students because they're working on their own level and thus are better prepared with foundational skills. A lot of students feel more accomplished because instruction is at their pace and level." Given this, several participants acknowledged that they might not see positive PARCC results in the first year because students and teachers were still adjusting to the model; because the TtO program was filling in gaps at first, over time student performance would increase. A few respondents commented that the gap

would be a challenge even in a traditional classroom setting. A couple of interviewees also underscored that the success of the program hinged on buy-in; "I think teachers and the administration need to believe in the philosophy that we're trying to meet students where they are. No matter what, we're not going to cram in testing."

The primary concern respondents raised was the tension between state accountability standards and the TtO model. Since state testing is used to determine whether teachers and principals are successful, they expressed concern that they were being evaluated unfairly in comparison with other schools and districts. Although TtO allows students who are below grade-level to obtain the skills they lack by meeting them where they are, state assessments largely address grade-level content. Therefore, the TtO program does not necessarily expose students to the grade-level content that they would experience in a traditional classroom. Despite the fact that students may be performing well on exit slips and the MAP assessment, these forms of assessments are not grade-level specific but rather assess students at their respective levels. Therefore, student performance on these tests may be promising, while student performance on statewide grade-level content assessments like PARCC may paint a different portrait.

Although participants across all schools stated that New Classrooms has been responsive and receptive to working with them to incorporate standardized test preparation, the test preparation was at most two months in duration, which some believed was not enough for students to be exposed to and master new content. Some individuals also remarked that adequately preparing students was a matter of preparing them not only with grade-level content but also with test-taking skills that TtO does not cover. Several respondents commented that this was toughest for eighth graders who have only had the opportunity to be in the program for one year and have not had the benefit of time to close the gap. A couple of individuals said that eighth graders were "in for a rude awakening" and "about to be really shaken" by the PARCC exam.

Of course, test prep can also have substantial downsides. Many teachers communicated that while New Classrooms accommodated their requests to incorporate test preparation in the weeks before the exam, the shift to test preparation was discouraging and frustrating for below-grade level students and resulted in increased behavioral issues. As one teacher put it, "Now we're stressing students out and they're feeling incompetent. They're seeing material they haven't seen…You're still setting up that gap between the kids who know and those who don't."

Another teacher added, "The past 6 months we spent building up their confidence, and they were actually learning. Now shifting to test prep defeats the whole purpose."

III. How is the TtO model working for students?

This section explores the extent to which staff felt TtO was effective, particularly for English Language Learners and special education students. This section also describes the level of student accountability and engagement under the TtO program, as well as the impact of the program on classroom culture and social environment.

English Language Learners

Virtually all interviewees expressed some level of concern that the TtO model did not fully address the needs of English Language Learners (ELLs) and students with disabilities (SwDs). Many staff asserted that TtO was designed only for students who had mastered a basic level of the English language and that the lack of resources in students' native languages created instructional difficulties, especially for recent immigrant students who arrived with no English language skills. Even with in-class support, schools felt stretched, given that one individual was required to support 10-12 students, who often possessed different levels of both English and mathematics proficiency. Several individuals explained that support staff did not have the capacity to reach all needy students during the short TtO sections. Of course, concerns about adequately addressing the needs of these students is not unique to the TtO format.[13]

Most schools instructed ELL students to use Google Translate. Again, this proved difficult for some recently arrived immigrant students who were not familiar with computers. Additionally, the time and effort it took to translate questions deducted from time devoted to student learning. Teachers felt that translating homework questions was particularly challenging because every student had a different homework packet. Furthermore, using Google Translate was an imperfect solution because translations were at times incomprehensible. In several schools, respondents described how teachers tried to work through these challenges by creating a repository of videos and alternative resources in Spanish, the most predominant native language in EPS. Teachers were uncertain, however, whether students were actually making use of these

---

[13] However, as noted above, analyses of MAP results found that TtO did not have differentially negative impacts on ELL and SPED students, accounting for their initial math ability. In contrast, we found significant negative effects of TtO for low-achieving students, regardless of their SPED or ELL status.

resources because they were not an official component of the TtO program. Moreover, it was at times difficult to find appropriate and comparable content in Spanish.

Despite the drawbacks of English-only content, a smaller handful of respondents argued that TtO's *lack* of language-specific content was actually improving student learning and building confidence among ELLs. Several participants explained that ELLs were becoming independent learners and taking greater initiative to translate and understand material on their own. As opposed to hand-delivering content in Spanish, the English-only TtO math content pushed students to learn English as well. Nonetheless, the majority of respondents requested Spanish-language content, including translated exit slips, homework, videos, and even possibly holding Live Investigations in Spanish.

Students with Disabilities

As with the instruction of ELLs, TtO's lack of content aimed specifically at students with disabilities (SwDs) created challenges for many teachers. The primary comment was on the lack of capacity to serve all SwDs. Respondents explained that in the past, in-class support staff usually served about three or four students in each class, but in the TtO model they now served as many as 12 students across a range of different grade levels in a 45-minute period. Therefore, support staff were forced to pick individual students to help, and students were not necessarily getting all the support services they were entitled to receive. Teachers of all modalities requested additional staff in order to accommodate so many students with different needs. Furthermore, teachers also asserted that some students required instruction below the content level that TtO offered.

However, individuals reported that like ELLs, SwDs were gaining self-confidence and independence, and were increasing their levels of classroom participation. These teachers claimed that in a traditional classroom setting they would have taught a specific curriculum regardless of whether students were prepared or not, but the TtO curriculum meets students where they are, offering students multiple opportunities to repeat lessons if needed. At the same time, the TtO model did not spoon-feed students, so they were also growing to be more self-sufficient. Suggestions to improve learning experience for SwDs include differentiation in TtO participation for students with mild versus moderate or severe learning disabilities, increasing the capacity of support staff, and adding lower grade-level content.

41

Student Accountability and Engagement

Interviews with teachers, school leaders, and New Classrooms staff elicited varied impressions of students' reactions to the pace and delivery of TtO instruction. Most respondents stated that students found the TtO program engaging and motivating. Several individuals remarked that it is difficult, if not impossible, for a traditional classroom setting to personalize instruction to students' diverse learning styles and needs, especially because they typically have a mandated lesson guide. These teachers valued the variety of modalities that the model introduces to teaching and learning; "It addresses the student who needs traditional classroom instruction from a teacher; it addresses the kinesthetic learner who needs to work on a project with a teammate; it addresses the visual learner who needs things drawn out. I feel like it draws on the multiple intelligences." Addressing differences in students' learning styles occurs not only through the different modalities, but also through access to different teachers. Unlike a traditional classroom setting where students are assigned one math teacher, in the TtO setting students have access to all TtO teachers. While one teacher might explain things in a certain way, another might take a different approach that works better for a particular student. Conversely, staff commented that some students found the lack of continuity challenging.

In terms of pacing, nearly all participants had positive responses to how the TtO model paced instruction to meet the needs of each student. They explained that struggling students no longer felt embarrassed because "they have the comfort of knowing [students] are all learning different things," and high-achieving students no longer felt held back. Several individuals commented that the personalized pace of instruction encouraged mastery of skills because students could repeatedly review content until they mastered it. Therefore, the model was preparing students with the skills they would need for high school, and students felt more self-assured and confident about their abilities.

In contrast, several respondents communicated that the TtO model was not for all students and that its success varied across students depending on their learning styles and motivation levels. These staff argued that some students who struggled with the new model missed having a connection with their classroom and building relationships with peers. However, a few of these respondents commented that that some of these students wanted one-on-one all the time, which was not realistic even in a traditional classroom setting; "Those that want their teacher all the time, whether they're in a regular setting or in here, they're still going to be lost."

42

Others commented on how student reactions and success were directly tied to students' intrinsic motivation. As one person described, "You have to have something inside you to say, 'I want to try.'" Students need to take initiative in order to succeed." Indeed, some teachers expressed that a considerable challenge was figuring out how to motivate students who did not have an innate sense of drive.

Furthermore, many teachers also stated that student perspectives of the TtO model varied by achievement level. While the TtO program encouraged greater self-confidence in struggling students, some high-achieving students were frustrated because they were used to sliding through, but now were challenged by the content—obviously a welcome outcome of TtO to the extent true. A number of interviewees discussed that high-achievers' parents were initially concerned when their children were not getting 100% on their work. In response, some schools held multiple meetings with parents to explain how the model worked. Although most students and teachers had learned to adapt to the new classroom environment, several individuals commented that the open space and noise-level were still distracting.

The overwhelming majority of respondents stated that they thought that the TtO program held students accountable for their learning. As one teacher put it, "Students have to take ownership, initiate learning, and ask questions. … (teachers) are just there to support their learning." Several participants noted that at the middle school level students should no longer be "babied" or "spoon-fed." As one individual described, "Nowadays kids are so used to having everything done for them. If we don't teach them how to be self-sufficient and independent from an early age, it's harder for them down the road, so that life skill is important." A few others agreed that while not all students were intrinsically self-motivated, students needed to learn in order to be successful, and that it was better learnt earlier than later. They further noted that students needed to know when to ask for help and how to become independent learners and thinkers. A few interviewees discussed the importance of increasing student accountability by having students turn in written class notes, showing their computer results, or designating specific roles and tasks for peer or group work.

Impact of TtO on the Learning Environment

Teachers reported that TtO had mixed effects on classroom management and the learning environment. Some teachers stated that TtO had positive effects on student behavior because it

provided built-in opportunities for physical movement. Others asserted that TtO prevented students from becoming frustrated by content that was too difficult or too easy. However, other teachers reported that the TtO model made classroom management more difficult because students have increased opportunities to communicate with one another and become distracted. In addition, many teachers believed that it was challenging for all teachers to set and maintain consistent expectations. For example, teachers stated that "students know that they can play some teachers but not others," and that "some teachers are more easy and I am more firm. It's hard because we're all trying to be on the same page." A small number of teachers reported that their teams were able to maintain consistent expectations for student behavior, but most teachers reported that this was a struggle.

Many teachers asserted that the TtO model made it more difficult to build relationships with students, which in turn negatively affected student classroom behavior. For example, one teacher reported that "behavior is controlled by the relationship you have with your students. Last year, I had better behavior and control because students would have gotten to know me quickly." While some teachers reported that they were eventually able to build student relationships within TtO, the majority reported that it had taken longer to build relationships than in previous years, and that the relationships that they had built were not as deep or effective. In contrast, a small number of teachers reported that students appreciated being able to build relationships with multiple teachers and that TtO was superior to being "stuck with a teacher that's not great."

Many teachers reported that TtO's "open" learning environment created noise and distractions that impeded learning, particularly for special education students or students with ADHD. Some teachers responded by creating pull-out spaces for special education students, including for the periods spent taking exit slips. Some teachers also reported that the TtO space is too small or "cram packed," which contributes to student distractions. A small number of teachers suggested that the TtO model would be "better" if it took place in traditional classrooms, with students transitioning from room to room.

Several teachers reported that classroom management was easier in live investigations than in computer-based modalities or group modalities. Teachers stated that students took advantage of computer-based modalities to "try to go to another website and watch Youtube

videos." One teacher stated that "Sometimes [SGCs or P2Ps] just become[s] a social gathering… they don't feel accountable for their work in those sessions."

IV. How do schools feel about New Classroom's support?

This final section details school feedback regarding the support that New Classrooms provides to school, both from infrastructure and staffing perspectives.

IT Challenges

Although a few respondents noted that being a one-to-one laptop school likely reduces potential challenges relating to the TtO culture and expectations, a couple of individuals remarked that the challenge was with getting teachers, not students, comfortable with technology, as students were already rather technologically savvy and adept in troubleshooting. Nonetheless, responses suggest that schools faced varying degrees of IT challenges with the TtO program. While older buildings tended to face more difficulties, respondents across all schools described regular network and bandwidth issues. Students and teachers often faced trouble logging into their portals and with portals that froze on them or kicked them out. This was a particular problem during exit slips, when the portal would shut down and sometimes grade the exit slips before students were able to finish. As the problem was most acute during exit slip time, some individuals conjectured that the system did not have the capacity to serve so many individuals at once.

New Classrooms Support Staff

The overwhelming majority of interview respondents reported that they were satisfied with the level and quality of New Classrooms support. Teachers and principals appreciated being able to choose their Math Director because it was central to ensuring teacher buy-in. They also commented that New Classrooms had encouraged dialogue with schools and had been extremely responsive to feedback. For example, when schools requested scheduling time to prepare for the PARCC assessment, New Classrooms worked with each school to devise a plan that worked for them. While teachers and principals stated that there has been a fair amount of professional development, many explained, "You can only learn by doing. It's one thing for them to show you, but you must do it yourself." Thus, schools have been learning as they go, and in this sense,

the level of support from the New Classrooms staff has been of utmost importance. Most individuals were appreciative of how New Classrooms' support and availability, and they underscored that New Classrooms was "one of the few programs that actually puts people on site and actually does something to support (schools) on an ongoing basis."

However, a small handful of respondents did not feel adequately supported by the New Classrooms staff. These individuals reported that while New Classrooms staff has been open to hearing their concerns and questions, they have not responded to the challenges and issues raised and have not offered impactful feedback on how to improve instruction within the program. One participant in particular felt that because certain on-site New Classrooms staff members had limited teaching experience, the support they provided was of limited value.

Interview responses highlighted a few recommendations for improving New Classrooms support. For one, some individuals remarked that it would better if New Classrooms entrusted schools with greater responsibility; "Schools would feel more comfortable if some responsibilities were shared from beginning. As a new program, you want to make everything perfect, but the result is that schools sometimes can feel helpless without support." Furthermore, others requested having more opportunities to share ideas and best practices across schools. For example, one teacher also suggested including a comments section on the portal so that teachers could directly upload comments and concerns rather than having to report them to the designated on-site New Classrooms staff.

CONCLUSIONS AND RECOMMENDATIONS

This study described the effects and implementation of *Teach to One: Math* during the first year of implementation in five schools in Elizabeth, New Jersey. Overall across all grades, our CITS models found no effect of *TtO* on student mathematics learning using state standardized assessments, but a modest negative effect in fifth grade. In contrast, our analyses of student scores from the Measures of Academic Progress identified an overall modest, negative effect on student mathematics learning between the fall and spring of year one. However, this negative effect occurred solely during the initial fall-to-winter implementation period: TtO and non-TtO students gained mathematics skills at comparable rates during the subsequent winter-to-spring period.

These findings suggest that TtO teachers and students may have experienced an initial adjustment period when student learning suffered. However, there is some evidence that as TtO schools became more familiar with the program during the latter half of the initial implementation year, student learning stabilized. Our qualitative findings suggest a steep learning curve on the part of teachers, supporting the notion of an initial adjustment period. The high levels of teacher buy-in combined with stabilized student learning outcomes in the second semester leave us optimistic about the potential for student growth in the coming years. Our evaluations of the year-2 and year-3 implementation periods will establish whether outcomes among TtO students continued on this upward trajectory.

A Fundamental Disruption

Unlike many educational reforms, TtO represents a fundamental disruption of traditional classroom process, and a distinctive approach to mathematics teaching and learning. In short, it asks a great deal of teachers and students. Given this, it is important to bear in mind that these results are from the first year of implementation. It is quite common for deep, meaningful interventions to find no (or even negative) impacts in their first year. This is true even for reforms that have been in place for a substantial period time: the approach is quite novel for those tasked with its implementation, even though it is quite familiar to its developers.

Regarding our conversations with school staff, we were both surprised and impressed by teachers' willingness and openness to the model, particularly with teachers who have been in the classroom, in some cases, for decades. Among virtually all teachers, support for the logic behind

TtO was broad. Many of the concerns they did raise about particular elements of TtO are well-known by New Classrooms staff, who have, according to these teachers, responded appropriately where possible. Other criticisms of the model may take further discussion and negotiation to fully resolve. We look forward to visiting with these same teachers again during the second year of TtO implementation in the Elizabeth Public Schools. The recommendations below are based largely on CPRE's qualitative analyses from year one:

Recommendations

1. **Increase time for teacher collaboration, common planning, and shared professional development.** In contrast to traditional, isolated, classroom instruction, the TtO model requires that teachers maintain shared and consistent expectations for student culture and behavior. In addition, the TtO model requires that all teachers (including those teaching ILZs) feel comfortable teaching a broader range of content than traditional models, including the full span of TtO grades served by the school. Shared PD on classroom management, pedagogical skills, and Grade 5-8 math content may help teachers adjust to these changes.

2. **Provide teachers planning time early in the day.** Some teachers reported that the short turnaround time between content/student assignments and instruction left them inadequate time to prepare. Scheduling TtO instruction for the middle and end of the day so that teachers can prep in the morning may provide adequate time to gather resources, prepare lessons, and plan differentiation for the unique needs of their students. Of course, doing so may require (complicated) adjustments to school-wide schedules.

3. **Evaluate the P2P and SGC modalities.** Classroom observations indicated that approximately half of students in P2P and SGC modalities choose to work independently rather than collaboratively. Revising tasks to *require* collaboration and teamwork could address this shortcoming, as could assigning roles within each group (e.g. facilitator, recorder, time-keeper, etc.). In addition, teachers suggested that these modalities would benefit from increased student accountability and fewer, more rigorous, questions.

4. **Continue to make LINs and Tasks more open-ended and rigorous**. Technological constraints necessitate that computer-based modalities focus on procedural fluency. This creates a corresponding need for LINs and Tasks to address the critical thinking and real-

world application elements of the Common Core. Consider scheduling double-blocks for Tasks and LINs to allow more time for teacher feedback and student discourse.

5. **Ensure there is a coherent and comprehensive rationale for the grading system, and that it is transparent to all teachers**. Many teachers expressed confusion or dissatisfaction with the TtO grading system, which they believe artificially inflates grades. Teachers, families, and students might benefit from opportunities to explore and discuss the relative characteristics and advantages of criterion-based vs. growth-based grading systems. Also consider exit slips that could allow for partial credit, multiple steps, or written explanations of student thinking.

6. **Continue to improve teachers' access to data and resources through the TtO portal**. Although teachers were generally pleased with the function of the TtO portal, some expressed a desire to view reports for groups of students instead of undertaking the time-consuming work of viewing each student individually. In addition, teachers sought the ability to observe and monitor how students were progressing through their individual skills maps, which might address teachers' feelings of separation from the learning process. Finally, teachers sought access to "teacher versions" for the P2P that were similar to the resources currently available for LINs.

7. **Increase supports for ELLs and students with disabilities**. Foreign-language content and assessments would obviously increase the ability of ELLs to access TtO content. Resources and strategies for serving students with disabilities would also help teachers provide high-quality supports for all students. In the absence of TtO supports, schools have attempted a variety of divergent strategies (e.g., pull-out spaces, dedicated SPED/TESOL teachers), but it is unclear to what degree these strategies are effectively supporting students.

8. **Address technical and bandwidth issues**. The student and teacher experience would be improved by continued improvements to the stability of the TtO platform and expansions of school-based internet bandwidth. The ability of New Classrooms to address these issues if obviously constrained in some respects.

9. **Provide additional opportunities for teacher-student feedback and relationship-building**. This could include providing more time in the MAs for reviewing homework, scheduling an "extra help session" within the school day to allow teachers to work with

students one-on-one, teacher-led activities across a double-block of time, or any other activity that could improve students' feelings of belonging and community.

References

Barr, R., & Dreeben, R. (1983). *How schools work*. Chicago: University of Chicago Press.

Bloom, H.S. (2003). Using "Short" Interrupted Time-Series Analysis To Measure The Impacts of Whole-School Reforms: With Applications to a Study of Accelerated Schools. *Evaluation Review*, Vol 27 No.1, pgs, 3-49.

Bruner, J. S. (1978). The role of dialogue in language acquisition. In A. Sinclair, R.J. Jarvelle, & W.J.M. Levelt (Eds.), *The Child's Concept of Language*, 241-256. New York: Springer-Verlag.

Cole, R., Kemple, J.J., & Segeritz, M.D. (2012). *Assessing the Early Impact of School of One: Evidence from Three School-Wide Pilots*. New York: Research Alliance for New York City Schools, New York University.

Corno, L. (2008). On teaching adaptively. *Educational Psychologist*, *43*(3), 161-173.

Lee, V.E., & Ready, D.D. (2009). The U.S. high school curriculum: Three phases of research and reform. *The Future of Children*, *19*(1), 135-156.

Martinez, J.F., Stecher, B., & Borko, H. (2009). Classroom assessment practices, teacher judgements, and student achievement in mathematics: Evidence from the ECLS. *Educational Assessment*, *14*(78), 78-102.

Newcomb, T. (2016). Teach to One: Inside the Personalized Learning Program that Bill Gates Calls the "Future of Math." Available at: https://www.the74million.org/article/teach-to-one-inside-the-personalized-learning-program-that-bill-gates-calls-the-future-of-math

Oakes, J. (1985). *Keeping Track: How Schools Structure Inequality*. New Haven, CT: Yale University Press.

Pallas, A. P., Entwisle, D. R., Alexander, K. L., & Stluka, M. F. (1994). Ability-group effects: Instructional, social, or institutional? *Sociology of Education*, *67*, 27-46.

Gates Foundation (2014). *Early Progress: Interim Research on Personalized Learning*. Seattle, WA: Author. Available at: http://k12education.gatesfoundation.org/student-success/personalized-learning/early-progress-interim-report-on-personalized-learning-rand/

Raudenbush, S. W. (2009). Adaptive centering with random effects: An alternate to the fixed effects model for studying time-varying treatments in school settings. *Education Finance and Policy*, *4*(4), 468-491.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and data*

analysis methods (2nd ed.). Thousand Oaks, CA: Sage Publications.

Raudenbush, S.W., Bryk, A.S., & Congdon, R. (2013). *Hierarchical Linear and Nonlinear Modeling (v. 7.01)*. Lincolnwood, IL: Scientific Software International, Inc.

Rockoff, J. (2015). *Evaluation Report on the School of One i3 Expansion*. New York, Columbia University Business School.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.

Snow, R.E. (1980). Aptitude, learner control, and adaptive instruction. *Educational Psychologist*, *15*, 151-158.

Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Wolf, M.A. (2010). *Innovate to Educate: System [Re]Design for Personalized Learning*. Washington, D.C.: Software & Information Industry Association.