

**The Achievement Network i3 Evaluation:
Year-One Survey Findings**

**The Center for Education Policy Research
at Harvard University**

November 8, 2012

Table of Contents

1. Introduction	3
1.1. Literature Review	5
1.2. ANet Model & Theory of Action	8
<i>1.2.1. Intervention Inputs</i>	8
<i>1.2.2. Leader & Teacher Actions</i>	9
<i>1.2.3. Student Outputs & Outcomes</i>	10
2. Research Design	12
2.1. Group Equivalence	13
2.2. Year-One Surveys	13
<i>2.2.1. Leader Analytic Sample</i>	14
<i>2.2.2. Teacher Analytic Sample</i>	14
2.3. Year-One Site Visits	16
3. Findings	17
3.1. Intervention Inputs	18
<i>3.1.1. Aligned Assessments</i>	19
<i>3.1.2. Logistical Support</i>	22
<i>3.1.3. Training and Coaching</i>	24
<i>3.1.4. Network Activities</i>	27
3.2. Leadership Actions	27
<i>3.2.1. Leadership Action: Establish Goals</i>	28
<i>3.2.2. Leadership Action: Analyze Results</i>	29
<i>3.2.3. Leadership Action: Facilitate Data Meetings</i>	30
<i>3.2.4. Leadership Action: Ensure Teacher Action</i>	30
3.3. Teacher Actions	31
<i>3.3.1. Teacher Action: Establish Goals</i>	32
<i>3.3.2. Teacher Action: Analyze Results</i>	32
<i>3.3.3. Teacher Action: Backward Planning, Action Plans & Reteaching</i>	34
3.4. Student Outputs & Additional Findings	36
<i>3.4.1. Student Outputs & Outcomes</i>	37
<i>3.4.2. Beliefs about Interim Assessments</i>	38
<i>3.4.3. Beliefs about Summative Assessments</i>	39
<i>3.4.4. School Culture</i>	39
4. Conclusion	41

5. References	44
Appendix A. Wave One Samples	46

1. Introduction

Efforts to promote the use of student performance data to inform instructional decisions are widespread in American education. Yet there is little systematic evidence on whether and under what conditions data-driven instructional programs impact teacher practice and, ultimately, student outcomes. We seek to address this shortcoming by studying the impact of an intensive data-driven instructional program, the Achievement Network (ANet), on educator behavior and student achievement in participating schools.

The Achievement Network was founded in 2005 to help schools improve mathematics and English language arts achievement in grades 3-8, particularly for high-need students. The ANet model combines high-quality assessments aligned to state standards, logistical support, educator coaching, and peer networks in order to help participating schools to make more effective use of data. In 2010, ANet received an Investing in Innovation (i3) grant to expand and evaluate its program and designated the Center for Education Policy Research at Harvard University as its evaluation partner. Our ongoing study employs a matched-pair, school-randomized design to estimate the impact of the ANet model on teacher and leader behavior, culture around data use, and student achievement in schools participating in the i3 expansion. Our primary outcome of interest is the level of student achievement after two years of ANet participation.

This report presents initial findings on ANet's impact on teacher and leader behavior based on data collected through surveys administered in all i3 evaluation schools and site visits to a subset of participating schools during the first year of program implementation. During year-one site visits to treatment-group schools, leaders and teachers consistently reported that prior to implementing the ANet model they had used interim assessments, analyzed student assessment

data, and more generally considered themselves to be “data-driven” organizations. Year-one survey results from control-group schools similarly indicate that “business as usual” in these schools includes the administration of interim assessments and efforts to use the information from these assessments to inform their instructional practice.¹

These patterns have important implications for the interpretation of both the year-one survey results presented in this report and the overall evaluation. In particular, survey results indicating the absence of an impact on specific leader or teacher behaviors ANet seeks to encourage does not necessarily indicate that this program component is not being implemented in participating schools; it may be that the component is commonly practiced in both ANet schools and those to which they are being compared. The results presented in this report are therefore most useful for identifying the most distinctive elements of the ANet model. Doing so is critical as our eventual estimates of the impact of the ANet model on student achievement will capture its effects over and above those of typical efforts to use interim assessment data to inform instruction in the types of schools in which ANet is currently expanding.

The report is organized as follows. The remainder of this introduction briefly reviews the extant literature on interim assessment and data-driven instruction and describes the logic model of the ANet intervention. In the second section, we provide an overview of the research design for the overall evaluation, specifics of the survey design central to this report, and a discussion of our analytic samples. The third section presents our main findings, while section four concludes. Appendix A provides a detailed description of the school sample, from recruited schools to the leader and teacher analysis samples.

¹ For example, all control school leaders and 92 percent of control school teachers responding to the survey reported that their school administered formal interim assessments during the 2011–12 school year.

1.1. Literature Review

Perie, Marion, and Gong define interim assessments as those “administered during instruction to evaluate students’ knowledge and skills relative to a specific set of academic goals in order to inform policymaker or educator decisions at the classroom, school, or district level” (2009: p. 6). A key feature of such assessments is that the results can be used by teachers to inform instructional decisions. The theory of action behind data-driven instruction is that providing teachers with interim assessment data and, in some cases, other supports (e.g., training or coaching) allows them to target instruction to areas where student mastery fell short and, therefore, improves student performance. However, there is limited research on the effect of interim assessment programs on teacher practices and student outcomes, and which supports may be needed to facilitate instructional improvement.

Several studies have examined whether, in fact, teachers change their instructional behavior in response to the availability of interim assessment results. A study of the use of interim assessment data in Providence Public Schools concludes that teachers aligned instruction to the assessments and used the assessment data to meet student needs (Clune & White, 2008). In a study of the implementation of No Child Left Behind in California, Georgia, and Pennsylvania, teachers reported that interim assessments were useful for “identifying and correcting gaps in curriculum and instruction” (Stecher et al., 2008: p. 66). However, both studies were based on teacher self-reports and did not investigate the specific ways in which teachers changed their practice or the supports necessary to bring about instructional improvement. In her 2010 review of research on interim assessments, Shepard concludes that:

More frequently, interim assessment results appeared to be used, item by item, to reteach the steps in the problems that were missed. In most cases, there was no

mention of attending to the underlying concept or to diagnosing more specifically where a student's understanding was breaking down" (2010: p. 255).

Indeed, in a recent study based on interview and classroom observation data for 32 teachers in two school districts, Riggan and Olah find that, "Interim assessments are most often used to identify weak content areas or students within a class" rather than to modify instruction (2011: p.11).

Two recent studies found no evidence of an effect of data-driven instruction on student achievement in traditional public schools. Studying a formative assessment program in reading, Quint, Sepanik, and Smith (2008) found no statistically significant improvement in the reading test scores of students whose teachers were exposed the program compared to students of teachers in comparison schools. Henderson, Petrosino, Guckenburg, and Hamilton (2007, 2008) similarly found no statistically significant impact of quarterly mathematics assessments on school mean mathematics achievement. Both of these studies employed nonexperimental designs, however, and the latter lacked sufficient statistical power to detect even modest positive effects.

In contrast, a recent study of New York City charter schools found that intensive use of data to inform instruction was among five practices associated with variation in schools' effectiveness in raising student in mathematics and English language arts (ELA) (Dobbie & Fryer, 2011). In particular, schools that administered five or more interim assessments during the school year and utilized four or more differentiation strategies reported math gains that were 0.08 standard deviations higher than comparison schools. Although based on observational comparisons, these results indicate that instructional data use may contribute to the success of some high-performing charter schools serving disadvantaged students.

Addressing the lack of experimental studies of data use and student achievement, Carlson, Borman, and Robinson (2011) examined the impact of whole-district data-driven reform initiative developed by the Center for Data-Driven Reform in Education (CDDRE). The intervention focused on the targeted use of interim assessment data and support from consultants in data interpretation. With assistance, district and school leaders were expected to use the results of the interim assessments and other available data to select and adopt appropriate evidence-based reforms. The intervention generated statistically significant improvement in mathematics achievement, but a lack of detailed implementation data left the researchers unable to examine how or why the improvement in mathematics achievement occurred and why similar gains were not evident in ELA. Dobbie and Fryer (2011), who also found a stronger relationship between instructional data use and achievement gains in math than in ELA, hypothesize that teachers may be less clear on how to use interim assessment data to inform ELA instruction compared to mathematics.

In sum, prior literature confirms the need for larger and more rigorous studies of the impact of data-driven instructional programs on student performance. Additionally, research highlighting changes in teacher behavior and instruction is needed to shed light on how and under what circumstances the availability of data from interim assessments can lead to changes in student performance. Our evaluation addresses both these shortcomings through an adequately sized sample making use of a matched-pair school-randomized design. This report focuses specifically on data from our year-one survey and site visits in order to document any immediate changes in teacher and leader behavior as a result of participating in the ANet program.

1.2. ANet Model & Theory of Action

As noted above, the ANet model combines high-quality standards-aligned assessments, logistical support, educator coaching, and peer networks to enable schools to make more effective use of data as a tool for increasing student achievement. This model is based on the belief that if teachers are provided with timely data on student performance from interim assessments tied to state standards, if they are taught how to use that data to identify student weaknesses, and if they have knowledge of how to improve the performance of students who are falling behind, then they will become more effective at identifying and addressing gaps in student learning. This will, in turn, improve student performance in their school, particularly for high-need students (see Exhibit 1).

1.2.1. Intervention Inputs

The key components of the ANet intervention model include:

1. Aligned Assessment: Four assessments administered every 6-8 weeks in mathematics and English Language Arts aligned to state content standards;
2. Logistical Support: Assessment-based performance reports identifying school, class, and student strengths and gaps that are available within 48 hours;
3. Training and Coaching: Coaching 4-14 times a school year and observations annually, along with two network-wide professional trainings annually; and
4. Network activities: Reports showing schools' performance relative to other network schools and two network meetings of member schools annually.

The core of the model is the administration of customized assessments in English language arts and mathematics that are tied to state standards and administered regularly throughout the school year (Aligned Assessment). Reports produced within two days of students taking the exam break

down the performance of classrooms and individual students. ANet provides a detailed protocol to guide each team's use of assessment data to assist in identifying gaps in meeting state standards (Logistical Support). School leaders receive regular coaching geared towards helping teachers learn how to embed data-driven decision making in everyday routines (Training and Coaching). The schools also participate in a network of similar schools in their region intended to foster both competition and the sharing of successful practices. Through these networks, teachers and leaders have the opportunity to learn from other teachers facing similar challenges (Network Activities).

1.2.2. Leader & Teacher Actions

The ANet model is a whole-school reform model designed to imbed data-driven decision-making in school leaders' and teachers' everyday practice. ANet believes that student interim assessment data is useless if educators do not know how to interpret it and address, through instructional interventions, any weaknesses it identifies. Therefore, there are several components of the logic model that school leaders and teachers are expected to demonstrate through their actions.

For school leaders, the first expectation is that they establish a Data Leadership Team typically consisting of the school principal, assistant principal (if applicable), and grade-level or content-area instructional leaders. The Data Leadership Team has six main purposes: to establish ambitious but feasible goals for student achievement, manage the administration of the assessment, analyze assessment results with the leadership team to identify trends and prioritize standards, facilitate data meetings with teachers, ensure teachers take instructional action and reflect on its impact, and use data to build a positive school-wide culture. ANet's approach to coaching aims to build leader capacity through a gradual release of responsibility model; the

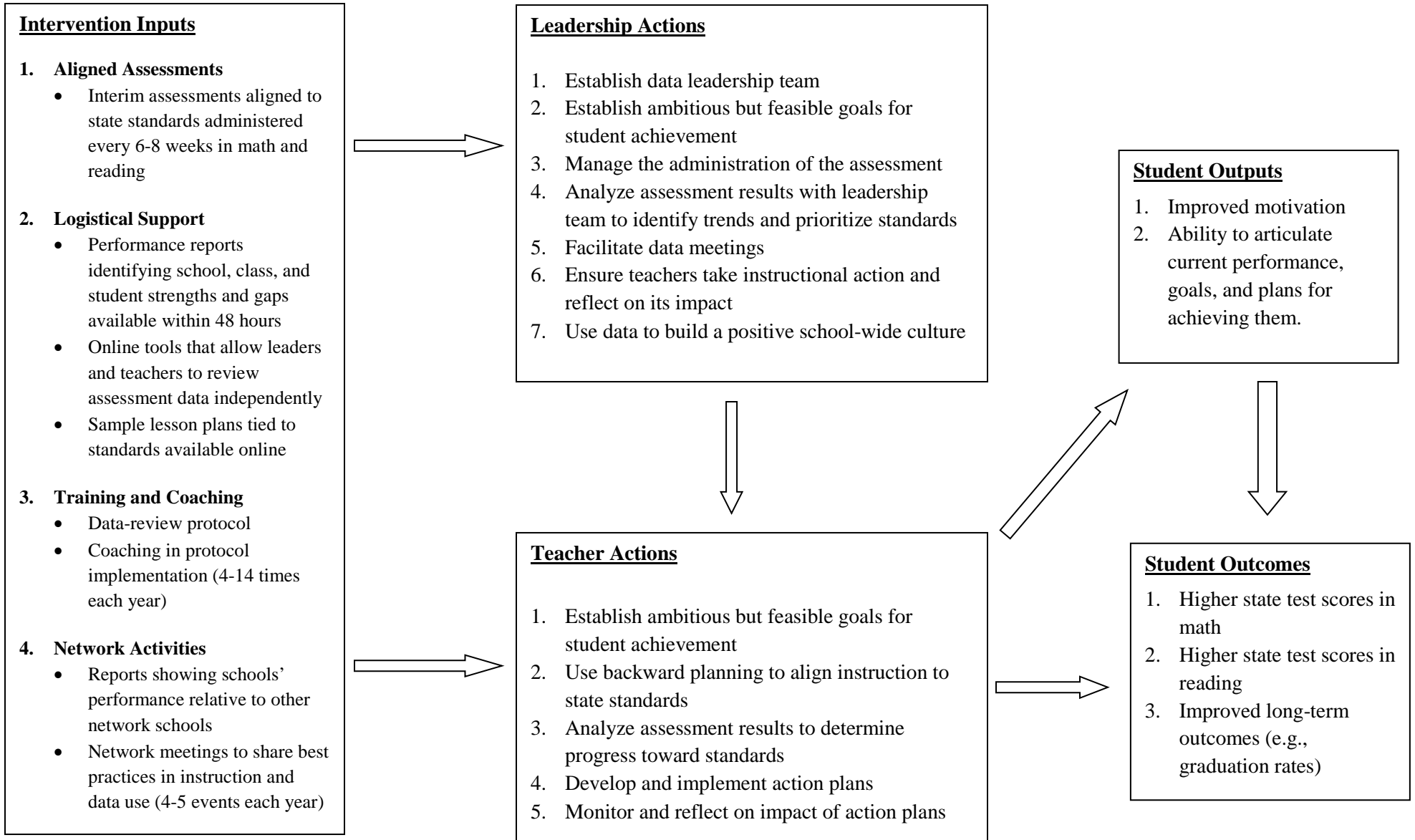
intent is for these actions to be modeled by the ANet coach, but eventually become the responsibility of the Data Leadership Team.

For their parts, teachers are expected to use backward planning to align their instruction to state content and performance standards. With this as a foundation, they are also expected to analyze assessment results to determine progress toward standards and use this information to develop and implement action plans that address students' gaps in knowledge. Once they have implemented these action plans, teachers are encouraged to assess and reflect on their impact as part of a continuous cycle of instructional improvement.

1.2.3. Student Outputs & Outcomes

Students in ANet schools are also expected to exhibit some outputs or intermediate outcomes. Through the sharing of interim assessment results, students are expected to exhibit greater motivation to learn, as well as to articulate their own performance goals and plans to achieve them. The primary outcome of interest for the ANet intervention is higher state summative assessment scores in math and reading. However, it is hoped that short-term impacts on test scores will translate into improvements in longer-term student outcomes, such as high school graduation and post-secondary success.

Exhibit 1. Logic Model for The Achievement Network



2. Research Design

To study the impact of the ANet program, we employ a matched pair school-randomized design that assigns schools to the treatment group or control group within pair and within district. Schools were matched on achievement and demographic variables, including prior year test scores, free and reduced-price lunch status, and racial/ethnic composition. One benefit of this design is that if and when attrition from the study sample occurs (due, for example, to a school closure), the relevant school and its pair can be dropped from the sample without undermining the internal validity of the results (Imai, King & Nall, 2009). Schools assigned to the treatment group are scheduled to receive the ANet intervention in the 2011-12 and 2012-13 school years. Schools assigned to the control group continued with “business as usual”, which may include some other type of data-driven instructional program. These schools will have the opportunity to implement the ANet model at a subsidized rate in the 2013-14 school year.

Seventy-one schools located in 5 districts (Chelsea Public Schools, Springfield Public Schools, Boston Public Schools, Jefferson Parish Public Schools, and the Chicago Public Schools) were recruited to participate in the study, including 42 elementary schools, 5 middle schools, and 24 k–8 schools. Although the racial/ethnic composition of the students enrolled in these schools varies across districts, all of the schools serve high proportions of students eligible for a subsidized lunch and have large shares of students who are not performing at proficient levels on state math and reading tests. These 71 schools were randomly assigned to treatment and control conditions, resulting in 36 treatment schools and 35 control schools included in our

teacher/leader impact sample.² Appendix A provides a detailed overview of the process for recruiting and randomizing schools and the definition of the analytic sample for this report.

2.1. Group Equivalence

Table 1 compares the demographic composition and prior achievement of the treatment and control group schools in our analytic sample. The first two columns present the means for the treatment and control groups. We also present the treatment coefficient and its corresponding p-value from OLS regressions that parallel those used to assess the impact of ANet treatment below.

The table confirms the absence of significant differences on demographic variables between treatment and control schools. Schools in both groups have a large percentage of free and reduced-price lunch students and have a majority of black and Hispanic students. Prior achievement is measured as school proficiency rates on state summative assessments two years prior to implementing the ANet model. The 6th grade ELA proficiency rates of treatment schools are lower than those of control group schools by a marginally significant amount. Because schools were matched on a transformed proficiency score that took into account the scores from all of the grades that ANet works with, this likely reflects chance variation in performance in this one grade.

[Table 1]

2.2. Year-One Surveys

Year-one school leader and teacher surveys were administered from May through July of 2012. The surveys provide information on attitudes and practices around data use, perceptions of

² Due to an odd number of recruited schools in one district, one “pair” of matched schools includes three schools, two of which were assigned to the treatment group.

interim assessments, and, in treatment schools, implementation challenges. Several considerations informed the development of survey items. First, we included items that would allow us to compare the extent to which educators in treatment and control group schools exhibit various actions specified in the ANet logic model. For purposes of measuring implementation fidelity, items were also designed to capture variation across treatment schools in model implementation. Finally, we undertook a review of the literature and related surveys on instructional data use to ensure that the surveys included items of relevance to the field.

2.2.1. Leader Analytic Sample

Of the 71 schools in the sample, we received responses from 57 school leaders representing 29 control schools and 28 treatment schools. Comparisons of treatment and control school characteristics and school leader demographics presented in Table 2 confirm that treatment and control respondents and their schools were comparable in terms of their observable characteristics. In both groups, a majority of leaders are female and identify as white or black. Nearly 90 percent hold a master's degree. Although the mean number of years of leadership experience is roughly 6 years, they have, on average, 18–20 years of total experience in their current district.

[Table 2]

2.2.2. Teacher Analytic Sample

From the 71 schools in the sample, a total of 654 year-one teacher surveys were received. The target population for the survey was generated from ANet rosters (treatment schools) and similar rosters obtained directly from control school leaders. In both cases, rosters were intended to include only those regular classroom teachers of grades 3 through 8 who taught mathematics

and/or English-language arts (i.e., the teacher sample). However, the rosters and, therefore, survey responses in practice contained some teachers who should not have been included in the teacher sample. We therefore created a variable to indicate which teachers were in-scope (i.e., part of the analytic sample), with the intention of ensuring that this sample corresponds to the set of teachers instructing students in subjects in which summative achievement data will be available.

In-scope teachers were defined as those who reported teaching at least one of grades 3 through 8 and they either 1) reported teaching ELA or reported some amount of ELA instructional time greater than zero or 2) they reported teaching math or reported some amount of math instructional time greater than zero or 3) they reported that they were an elementary/general education teacher or some amount of math or ELA instructional time greater than zero. This method was the least restrictive of several strategies of defining in-scope teachers, but did exclude about 125 teacher survey responses.³ The final analytic sample therefore consists of 527 teachers in 66 schools.

Table 3 confirms that in-scope treatment and control school teachers were equivalent on all available demographic characteristics.⁴ In both groups, a majority of teachers are female and identify as White. About two-thirds of teachers hold a master's degree. The mean number of years of teaching experience is reported at about 12 to 13 years, with about 7 years of experience in their current schools.

[Table 3]

³ Excluded respondents for whom an assignment could be determined generally reported teaching a non-standard ANet grade (e.g., grade 2) or subject (e.g., science or social studies), or they reported that they served in a coach or specialist role (e.g., ELL or special education teacher). There was no difference in the proportion of teachers excluded from the treatment and control groups.

⁴ Grade level and subject area not tested.

Because response rates varied by treatment and control schools, we also examined whether there were any significant differences in achievement and school demographics and student performance between the treatment and control schools. Tables 4 and 5 present results from this analysis. The results in these two tables do not point to any observable differences in school demographics between the survey respondents from the treatment and control schools, suggesting that at least on observable dimensions voluntary survey response did not unduly affect the original matched-pair randomization.

[Tables 4 & 5]

2.3. Year-One Site Visits

During the third data cycle (Jan–March 2012), CEPR visited three schools in each of the four participating i3 networks (Chicago, Jefferson Parish, and Eastern and Western Massachusetts). The purpose of these site visits was to gather rich descriptive data on the implementation of the ANet model and the early impact on teacher and leader practices and school culture, as perceived by educators in those schools. Schools were selected to provide variation within each network in the degree of both ANet coach involvement and school adoption of the activities and practices specified in the intervention logic model.

Prior to visits, CEPR developed interview and observation protocols for each component of the site visit to standardize the data collection process from each of the 12 visited schools. Researchers collected data on the provision of resources and support from ANet, the structure and content of data meetings, and any adaptations of the ANet model to specific school contexts. They also sought to identify factors that may hinder or facilitate implementation of the ANet model.

3. Findings

Our primary analyses compare the responses to survey items asked of school leaders and teachers in both treatment and control schools. Given random assignment, statistically significant differences in these responses can be interpreted as the causal impact of implementing the ANet model in its first year. The findings are organized according to the intervention logic model, beginning with leader- and teacher-reports concerning each of the four key ANet program inputs: aligned assessments, logistical support, training and coaching, and network activities. We next turn to results pertaining to leader and teacher actions. In some cases, our survey data were insufficiently detailed to report on certain leader actions specified in the logic model. Additionally, some questions concerning leader and teacher actions that are highly specific to the ANet model were asked of the treatment group only; for example, only treatment school leaders were asked about their facilitation of data meetings. A final section presents findings on student outcomes and other relevant issues that are not formally part of the logic model. Exhibit 2 details each logic model component and the data source(s) on which findings concerning each component are based.

[Exhibit 2]

Results concerning ANet’s causal impact are based on OLS or Probit regressions of each survey variable on treatment status.⁵ We use two approaches to report the magnitude of effects: effect sizes (for Likert-scaled outcomes) and marginal effects (for dichotomous outcomes). Effect sizes (*d*) measure the magnitude of estimated treatment effects relative to the underlying variation in the relevant outcomes. In this report, we calculate effect sizes by dividing the

⁵ Because site visits to treatment schools suggested considerable variation in ANet implementation across networks, these regressions also include a set of network indicator variables in order to increase statistical power. An additional indicator variable for the three-school “pair” accounts for the differential probability of assignment to the treatment condition for these schools.

estimated treatment effect by the standard deviation of the relevant outcome among control group observations. A widely used benchmark for interpreting effect sizes suggests that effect-size estimates in a range around 0.2 are “small,” estimates around 0.5 are “medium,” and estimates in a range around 0.8 are “large” effects. However, effect sizes are relative and should also be considered within the context of the relationship between variables being examined. For example, a small effect on student achievement can actually translate to a substantively meaningful improvement in learning, but a small effect on some aspect of teacher practice (e.g., data use) may not be as meaningful if its connection with student outcomes is more tenuous.

Marginal effects (Pr) reported for dichotomous outcomes are produced based on a probit regression of each of these survey variables on treatment status. Marginal effects provide an estimate of the expected change in the outcome variable for each unit change in the independent variable. In the analyses presented here, they represent the difference between the treatment group and the control group members in the probability of responding “yes” versus “no” on a yes/no survey question. For example, an estimated marginal effect of 0.25 for a yes/no item asking about a particular practice would indicate that treatment group members were 25 percent percentage points more likely to report that they engaged in that practice than were control group members.

3.1. Intervention Inputs

The ANet model of data-driven instruction focuses on the provision of four main inputs: aligned assessments, logistical support, training and coaching, and network activities. One of the earliest indications that the program may succeed would be positive responses by school leaders and teachers in treatment schools regarding the presence and quality of these inputs. In fact, the

overall pattern of findings shows that educators from treatment schools rated several ANet program inputs higher than control-group educators rated similar inputs in their schools. These differences were especially notable with respect to the rigor of the ANet assessments and training provided by ANet coaches.

3.1.1. Aligned Assessments

School Leaders. All school leaders reported that their school administered formal interim assessments during the 2011-12 school year.⁶ The survey included a series of items asking leaders to rate the alignment of these interim assessments to their school or district curriculum, school or district curricular scope and sequence, state standards, and state end-of-year tests in the same subject. It also asked them to compare the rigor of interim assessments in each subject to that of their curriculum, state standards, and state end-of-year tests. While treatment-school leaders were asked to report on the alignment and rigor of the ANet interim assessments specifically, control-school leaders were asked to report on any interim assessments in use in their school (and, therefore, may have considered multiple interim assessment programs when responding).

Compared to their control-school counterparts, leaders in treatment schools reported that their ELA and math assessments were better aligned to and more rigorous than the state content standards, state end-of-year assessment, school or district curriculum, and school or district curricular scope and sequence (Table 6). The differences were especially pronounced with respect to rigor, with treatment-school leaders generally rating the ANet interim assessments a full point higher on a 1–5 scale that ranges from 1=“much less rigorous” to 5=“much more

⁶ All control-school leaders reported administering at least one interim assessment in math. All except one leader reported administering at least one interim assessment in English-language arts, reading, and/or writing.

rigorous” ($1.24 < d < 2.0$). Just one exception was reported: no statistically significant difference was found in the leader-reported alignment of the math interim assessment with their curricular scope and sequence ($d = .34, p = 0.22$).

[Table 6]

Teachers. Consistent with the leader survey, nearly all control-school teachers reported that they administered formal interim assessments (92 percent). The slight discrepancy between school leader and teacher responses on this item might be because these teachers are assigned to a grade or subject in which interim assessments were not administered or from confusion over the definition of “formal interim assessments.”⁷

Teachers in both groups rated themselves as having a “good” knowledge of the ELA and math standards, and agreed that the standards were of “good” quality. With this as context, teachers were asked the same series of alignment and rigor questions regarding their interim assessments. Compared to their counterparts in control schools, teachers in treatment schools reported that both the ELA and math assessments were more rigorous than the district or school curriculum, state standards, and state end-of-year test (Table 6). Patterns in teacher responses regarding the alignment of interim assessments differed from school leaders, however. Treatment school teachers reported that the ANet math interim assessments were not as well aligned to the school or district curriculum and curricular scope or sequence as were the interim assessments in use in control schools. It may be the case that math interim assessments in use in control schools are no better aligned than ANet assessments, but that teachers working with ANet became aware of alignment issues as a result of analyzing the interim assessment data more extensively. No

⁷ Since the teacher analytic sample is not a subset of the school leader analytic sample, it was possible that the discrepancy was because teachers who responded “no” were from schools for which there was no school leader survey and that the school leader would have also responded “no” had they completed the survey. A check of the data showed that this was not the case. In-scope teachers who responded “no” were often from schools with a responding principal and, in some cases, other teachers in the same school reported that interim assessments were administered.

statistically significant difference was found with respect to teacher reports of the alignment of math interim assessments to the state content standards or end-of-year assessment, nor with respect to the overall alignment of ELA interims (Table 6).

The results generally parallel findings from the year-one site visits, where teachers consistently commented on the rigor of ANet interim assessments while occasionally noting challenges in terms of alignment with their school's curriculum. For example, a Springfield teacher in an ANet school noted that the ANet assessments "[are] more rigorous, more aligned to MCAS, more resembling what an MCAS assessment would be." One Boston teacher, meanwhile, explained that she:

"went to the summer training and they talked about how...with the exception of 8th grade Algebra they had aligned with our curriculum completely and that is not the case at all.... So there was a lot of playing around with the standards and seeing where we could match things up and where things didn't quite match up...."

Additional survey items (not shown) provide more insight on how the use of interim assessment data differed between treatment and control schools. For example, teachers in treatment schools were 35 percentage points more likely to report that classroom-level interim assessment results were made public in their school ($p = 0.00$). This difference is likely attributable to participation in data meetings where results are shared, an ANet program component that may not be as common in other interim assessment programs.⁸ In addition, only about one-third of control-school teachers said that they had access to interim assessment results

⁸ This item was also asked of school leaders. Compared to their control-school counterparts, leaders in treatment schools were 21 percent points more likely to report that classroom level ANet results were made public in their school ($p = 0.13$). The smaller (and marginally insignificant) effect on leader responses may indicate that, from their perspective, student results were already public (prior to ANet) and are more comparable to control-group practices.

from other schools with similar students; such results are a central component of the data reported back to all teachers in ANet schools.

3.1.2. Logistical Support

Treatment School Leaders. Questions on this aspect of the logic model included several questions asked only of treatment leaders. Treatment leaders were asked about their knowledge and use of MyANet. Overall, 96 percent of leaders were aware of MyANet with the mean response regarding the use of MyANet’s individual features falling between “once or twice a year” and “monthly.” The various score reporting features of MyANet were among the most frequently used features and the quiz generator and Better Lessons used least often. In terms of the features’ usefulness, the mean response was typically between ‘somewhat useful’ and ‘very useful’ (Table 7).

[Table 7]

Treatment Teachers. Treatment school teachers were also asked about their knowledge and use of the MyANet site. Like school leaders, almost all of the treatment teachers were aware of MyANet (97 percent). The mean response regarding the use of MyANet’s individual features typically fell between “once or twice a year” and “once or twice a month.” Teachers most often used resources that supported their review of student assessment scores. Least often used were resources for comparing their students’ results with that of others within their network and Better Lessons. In terms of the features’ usefulness, the mean response was typically between “somewhat useful” and “very useful.” The pattern of responses regarding usefulness mirrored that of actual use; the more teachers used a particular feature of MyANet, the more useful they tended to rate that feature (Table 7).

Reports of the usage of MyANet are a bit surprising given the findings from site visit interviews during the third assessment cycle. Site visit data seemed to indicate that many teachers had not heard of or used MyANet with any frequency. It is possible that teachers' awareness of MyANet increased in the fourth data cycle; however, given the web-based platform for administering the survey, it is also worth considering that the likelihood of responding to survey is correlated with MyANet usage.

Treatment and Control Teachers. On the common items, treatment teachers reported higher satisfaction with timeliness ($d = 0.39, p = 0.00$; not shown) and clarity of the interim assessment data they receive ($d = 0.27, p = 0.03$; not shown). Research has demonstrated that the timeliness of data can be a significant roadblock in the effective use of interim assessment data for improving instruction (Kerr et al., 2006). Satisfaction with timeliness of the delivery of the ANet interim assessment data was something that came up often in site visit interviews. One Chicago teacher remarked, "I get the data right away...that makes it worth it." Another teacher from Boston said that this was "the first time for (her) that the data has been useful."

Control School Leaders. Control leaders were asked about the use of a protocol to guide data use. In total, 74 percent of leaders reported using a process or protocol to review interim assessment data (not shown). In each case, the majority of control-school leaders reported that their process or protocol included each of the following:

- evaluating how many students overall are on track to meet state standards,
- identifying which individual students are on track to meet state standards,
- identifying which specific standards many students have not yet mastered,
- identifying which specific standards each individual student has not yet mastered,
- analyzing incorrect student responses to identify common misunderstandings,

- communicating to teachers the specific students or skills on which they should focus their instruction,
- developing re-teaching plans.

Protocols used in control schools were most likely to include support for identifying standards on which many students failed to demonstrate mastery, analyzing incorrect student responses to identify common misunderstandings, and clarifying the specific students or standards on which teachers should focus their reteaching (85 percent). The protocols were least likely to include support for developing reteaching plans (65 percent). At the same time, site visit data indicated that many treatment school leaders and teachers felt that prior to ANet, they (in the words of a Chicago Data Leader) “were really missing out on the re-teaching piece.”

3.1.3. Training and Coaching

School Leaders. For most of the items regarding training and support, treatment leaders were asked about ANet support and control leaders were asked about general support for the use of interim assessment data. On the single common item, treatment and control leaders responded similarly to their satisfaction with support from the school or district for using interim assessment data (mean response between “would have liked a little more” and “adequate”; Table 8).

Among the items comparing support for interim assessment data use provided by the ANet to support from an external provider, treatment leaders reported far higher overall satisfaction ($d = 1.00, p = 0.00$). When asked about their satisfaction with specific training components (i.e., analysis, goal setting, instructional practice, and timeliness of support), the average response across all leaders was generally neutral to satisfied; however, leaders in

treatment schools reported significantly higher satisfaction ($0.94 < d < 1.29, p = 0.00$), with treatment-school leaders generally rating the ANet coach a full point higher on a 1–5 scale (Table 8). These findings are consistent with site visit data: School leaders and teachers consistently reported that their ANet coaches were responsive and supportive, and that working with them through the first three assessment cycles had been a positive experience.

[Table 8]

Teachers. Treatment and control teachers responded similarly to their satisfaction with resources to meet student needs and track their progress. However, treatment teachers reported higher satisfaction with support from the school or district for using interim assessment data ($d = 0.62, p = 0.00$; Table 8). This may indicate real satisfaction with district-provided support or it may be somewhat confounded with ANet support if teachers view ANet as a district-provided resource. In comparing support for the use of interim assessment data provided by ANet (or an external provider in control schools), treatment teachers were more satisfied ($d = 0.26, p = 0.03$). This is consistent with the positive feedback on ANet coaches collected during the year one site visits and illustrated by one teacher from Boston who remarked, “I couldn’t speak higher volumes of [our ANet coach]. She is so helpful and I would just love to bounce ideas off of her and ask for more help.”

When asked about specific support or training components, teachers in treatment schools reported significantly higher satisfaction with support for the analysis of assessment data ($d = 0.47, p = 0.00$), setting performance goals ($d = 0.39, p = 0.01$), improving instructional practices ($d = 0.36, p = 0.01$), and the response time for questions (of coach or training provider) ($d = 0.51, p = 0.00$; Table 8).

In addition to these specific items on interim assessment coaching and support, all teachers were asked to respond to a series of items on professional development (PD). Overall, agreement was highest for items asking about the correspondence of PD to the school improvement plan and whether PD helped build skills to meet student needs, and lowest on items about the inclusion of opportunities to work with teacher in other schools and follow-up by their own school leaders. Compared to their counterparts in control schools, treatment teachers reported higher agreement with items asking about building new skills to meet student needs ($d = 0.18, p = 0.07$), receiving follow-up support from leaders on implementing what was learned ($d = 0.28, p = 0.05$), and opportunities to work productively with colleagues in their school ($d = 0.26, p = 0.02$; Table 9).

[Table 9]

Unlike other items for which no statistically significant differences were observed, these aspects correspond to key components of the ANet model. Specific to collaboration, teachers explained that since the introduction of ANet, discussions around data with colleagues had become richer, more frequent, and more focused:

“Well once we’ve done the re-teaching portion and we sit down with our reflection meeting, we’re really coming back with our data – here’s where our kids started, here’s where they ended after the re-teach. We think ... this is an opportunity for us to share strategies and techniques, and what worked in the classroom, and I did this differently, and this is how it benefited us, so that’s ... really really been a helpful key for us to just to be able to collaborate in the reflection meeting and share techniques and strategies” (Chicago teacher)

3.1.4. Network Activities

School Leaders. Across all leaders, meetings with school leaders from within the district were typically held quarterly to monthly, while meetings with school leaders from other districts were rarer. Overall, schools leaders reported that these meetings were “somewhat” to “considerably” useful (not shown). When asked about specific discussions that were held during these meetings, only one significant difference was detected between treatment and control leaders: Treatment leaders reported that they discussed student achievement data in meetings with other school leaders less frequently than did their control school counterparts ($d = -0.60$, $p = 0.03$; not shown), a puzzling finding given our observations of ANet network meetings. Otherwise, leaders in treatment and control schools reported discussions of instructional improvement and school improvement in meetings with other school leaders with similar frequency.

3.2. Leadership Actions

Prior research shows that principal leadership can influence both the successfulness of schools reform in general and the degree to which teachers alter their practices to improve student achievement specifically (Johnson, Berg & Donaldson, 2005; Datnow, Park, & Wohlstetter, 2007). ANet recognizes the importance of school leaders in establishing both a positive message and effective practices around data-driven instruction. Therefore, treatment-school leaders are responsible for seven key actions that in turn should drive teacher actions and better outcomes for students. Several of these actions are relatively straightforward; establishing a data leadership team and managing assessment administration. However, others require substantial time and energy to implement; for example, analyzing assessment results and ensuring their teachers take instructional action. Overall, the results show that treatment school

leaders exhibit greater confidence in using interim assessment data for several key practices such as monitoring school performance over time. Teachers in treatment schools are also more likely to report that their school leaders encourage data use and provide feedback on instructional plans.

3.2.1. Leadership Action: Establish Goals

School Leaders. Leaders were asked to report whether and in which months they set goals for student performance. These items were recoded into two summary measures: the earliest time that goals were set and how many times they set goals throughout the school year. Nearly all (93 percent) school leaders reported that they had set student performance goals for the 2011–12 state test (not shown). On average, goals were set twice throughout the school year, with goals first set in September (Table 10).

As expected, a great percentage of school leaders in both treatment and control schools set school-wide goals (85 percent) with fewer reporting that they set goals for specific students (50 percent). Current achievement levels and statewide targets were most frequently reported to be the helpful for setting these goals (85 percent and 81 percent, respectively). Comparable data from other schools was least often reported as helpful in setting student performance goals (15 percent; Table 10).

[Table 10]

The only statistically significant difference between treatment and control leaders was found when comparing their level of confidence using interim assessment data to set schoolwide goals with treatment leaders reporting higher confidence ($d = 0.58$, $p = 0.03$; Table 10).

Teachers. Teachers overall gave high ratings to their school leaders' ability to communicate a vision for teaching and learning, set grade/classroom level instructional goals,

and set high standards for learning. Mean responses indicated ratings between “fair” and “good.” All teachers reported a similar rating of their school leaders’ ability to track student progress (not shown).

3.2.2. Leadership Action: Analyze Results

School Leaders. Among items coded under analysis of assessment results, school leaders reported with similar frequency using interim assessment to evaluate programs and teachers, compare performance of classes and groups of students, and refer students for services. Responses were similar between treatment and control leaders with the majority reporting that interim assessment data were used to set schoolwide and classroom-level achievement goals and to track student performance.

The one exception was that leaders in treatment schools were 42 percentage points more likely to use interim assessment data to compare their school’s performance with that of others ($p = 0.00$; Table 11). This may be due in part to the fact that only 39 percent of control school leaders reported that they had access to comparable interim assessment data from other schools serving similar populations of students. This may also be why treatment school leaders reported higher confidence in using interim assessment data to compare their school’s performance to that of others ($d = 0.61$, $p = 0.02$; Table 11). They also reported higher confidence in using interim assessment data to examine trends in schools performance over time ($d = 0.73$, $p = 0.01$), and identify students in need of instructional support ($d = 0.48$, $p = 0.06$). There was no difference in leaders’ reported confidence in the use of interim assessment data for evaluating the effectiveness of individual teachers or instructional programs.

[Table 11]

Teachers. Teachers from treatment schools were 19 percentage points more likely to report that the assistant principal led the work of analyzing interim assessment data as compared to their counterparts in control schools ($p = 0.03$; not shown). This could be a result of the assistant principal serving an official role on the ANet data team.

3.2.3. Leadership Action: Facilitate Data Meetings

Treatment School Leaders. Only one treatment leader survey item was coded under this element of the logic model. Overall, treatment leaders reported that they felt “quite confident” to “highly confident” leading a data meeting without support from their ANet coach. Among other items related to data meetings (treatment only), treatment leaders reported that the various components of the ANet data meetings were “somewhat” to “very useful,” with the means indicating more responses of “very useful” (Table 12).

[Table 12]

3.2.4. Leadership Action: Ensure Teacher Action

School Leaders. Leaders in all schools reported with similar frequency that they observed classroom instruction, participated in teacher team meetings, and developed individual teacher professional development plans. However, compared to control school leaders, treatment school leaders reported that they more frequently encouraged teachers to reflect on their instructional practice ($d = 0.55$; $p = 0.03$; Table 13). This result is consistent with year-one site visits, where many treatment school leaders (and teachers) explained that the added value of the ANet model was the focus on reteaching plans and reflecting on practice. One Chicago principal explained that:

“The foundation... was already set at (my school) around data and data practices. And the part of the cycle we had problems with was the re-teaching, that’s what we worked on effectively, and as I sat and watched someone, oh yeah that was our problem, we really don’t do that part right now. So ANET helped us get our acts together on that.”

Treatment leaders reported more frequently than control leaders that the person or people responsible for reviewing interim assessment data also observed teacher instruction ($d = 0.43, p = 0.09$), met with teachers to discuss data ($d = 0.54, p = 0.06$), and coached teachers in using data to monitor student progress ($d = 0.55, p = 0.02$) and develop reteaching plans ($d = 0.38, p = 0.10$). Leaders in treatment and control schools responded similarly on all other aspects of providing data support (Table 13). They also responded similarly in regards to their own confidence in leading teachers in analyzing interim assessment data (not shown).

[Table 13]

Teachers. Consistent with several of the key leadership outcomes of the ANet model, teachers in treatment schools gave higher ratings to their school leaders’ abilities to monitor the quality of teaching ($d = 0.21, p = 0.09$) and provide feedback on instructional plans ($d = 0.27, p = 0.04$). Teachers in treatment schools also gave higher ratings to their school leaders’ abilities to institute practices to encourage data use ($d = 0.22, p = 0.1$; Table 13).

3.3. Teacher Actions

Though quality program inputs and strong school leadership matters, much of the responsibility for realizing student achievement gains through the use of assessment data falls on classroom teachers. The ANet logic model specifies five key teacher actions. These are not one-

time actions; they require a shift in instructional practices toward a cycle of goal-setting, assessment, instruction, and reflection. Compared to their counterparts in control schools, the following results show that teachers in treatment schools are more likely to use data during common planning time to identify what should be retaught and when, and whether to reteach to the whole class, small groups, or individual students. They are also more likely to use common planning time to discuss lesson plans and instructional strategies.

3.3.1. Teacher Action: Establish Goals

Teacher-Reported. The majority of teachers (about 88 percent) reported setting goals for their students' performance on the state test and that goals were set an average of 2 to 3 times over the school year. Overall, they first set them early in the year; however, treatment teachers were less likely to report setting them in September ($d = -0.16, p = 0.05$) – and more likely to set them in November ($d = 0.18, p = 0.06$) and March ($d = 0.17, p = 0.05$) – when compared to control school teachers. This could be a product of the timing of the ANet interim assessment cycle; with the first and third cycles occurring around these latter two dates. Both groups were equally likely to set grade- and class-level goals, but treatment teachers were less likely to report setting individual student goals ($d = -0.20, p = 0.05$).

3.3.2. Teacher Action: Analyze Results

Teacher Reported. The frequency of teacher-reported collegial conversations or conversations during common planning time regarding data, student progress, student test results, and particular students who are not meeting standards was similar across both treatment and control teachers. Treatment teachers reported a greater frequency of reviewing interim

assessment results independently ($d = 0.22, p = 0.09$), with their team ($d = 0.36, p = 0.00$), with other teachers in their school ($d = 0.35, p = 0.01$), and with their school leaders ($d = 0.36, p = 0.01$; Table 14).

[Table 14]

In a series of items about the use of data during common planning time, teachers' mean responses were moderate to high for all items (Table 14). However, treatment teachers reported higher frequency of using data (of any type) during common planning time to identify what and when to reteach ($d = 0.43, p = 0.00$), understand student groupings when re-teaching a skill ($d = 0.38, p = 0.00$), and identify material to supplement the school curriculum ($d = 0.20, p = 0.04$). Treatment teachers also reported greater frequency in using *interim assessment data* to plan instruction ($d = 0.29, p = 0.00$) and monitor student progress ($d = 0.23, p = 0.04$). However, teachers in both the treatment and control groups reported similar frequency in the use of interim assessment data to adjust plans to meet student needs and identify lesson that were mastered. These response patterns may be indicative of the areas in which the ANet treatment differs from the other data-driven practices in use in control schools. Using data independently to adjust instruction to target standards that weren't mastered may be standard procedure in various data-driven instructional models; however, the value of ANet may be in its focus on teacher collaboration, and support for planning instruction and monitoring student progress.

Both groups reported similar confidence in their ability to identify individual students' mastery of standards (mean response indicating "confident"). Among treatment school teachers, the consensus was that reviewing class-level and individual student ANet interim assessment data, as well as using the data to identify a priority skill/standard, was a "somewhat" to "very useful" component of ANet data meetings (not shown).

Leader Reported. Leaders in treatment and control schools responded similarly in regards to the number of teachers who used data to identify students who need instructional support. However, compared to school leaders in control schools, treatment school leaders reported that more of their teachers used data to identify which standards students had not yet mastered ($d = 0.57, p = 0.02$; Table 15). This may be attributable to the specificity of the ANet protocols (compared to other programs' protocols) for identifying the key takeaways from their students' results (i.e., standard and subskills of focus); the step required prior to the development of reteaching plans.

[Table 15]

3.3.3. *Teacher Action: Backward Planning, Action Plans & Reteaching*

Teacher Reported. Across both groups, teachers more frequently had conversations about instructional strategies with teachers in their own schools. It was less common that they had similar opportunities to collaborate with teachers in other schools or districts. However, teachers in treatment schools reported modestly higher agreement that opportunities to discuss instructional strategies with teachers from other schools in their district were useful ($d = 0.17, p = 0.11$; Table 16).⁹

Compared to their counterparts in control schools, treatment school teachers reported a greater frequency of having discussions outside of and during common planning time about lesson plans¹⁰ ($d = 0.20, p = 0.07$; $d = 0.31, p = 0.01$, respectively) as well as instructional methods or pedagogy ($d = 0.21, p = 0.03$). Treatment teachers also reported stronger agreement

⁹ Probit model, $p = 0.09$.

¹⁰ These results refer to two similarly worded items. One item asks about "new" lesson plans, where the other does not. The relationship is stronger for the more generally worded item, possibly indicating that discussions about lesson plans were not always about the development of entirely new lessons.

that common planning time discussion influenced their own classroom practices as compared to their control school counterparts ($d = 0.18$, $p = 0.06$; Table 16).

[Table 16]

Teachers in treatment and control groups were equally confident in their ability to plan and modify instruction to meet students' needs. The mean response indicates that, on average, teachers responded "confident" to "highly confident" (not shown). Teachers in both treatment and control groups reported similar frequency of beginning their instructional planning by identifying the skill or goal, or state standard that students should master. The mean response indicates that teachers in both groups backward plan from target skills and standards "often" to "almost always." Teachers in both groups also responded similarly (but with decreasing frequency) to items regarding whether their planning included: creating differentiated instructional plans, using the curricular scope and sequence, and scheduling re-teaching within and outside their regular lesson plans or instructional time (Table 16).

In terms of the frequency with which teachers reported using grouping strategies during their re-teaching, teachers in both treatment and control schools reported with similar frequency that they re-taught to small groups and individual students; however, treatment school teachers were more likely to report whole-class re-teaching ($d = .29$, $p = 0.00$; Table 16). It may be the case that the ANet interim assessments more frequently identified standards that most or all students in a class had not yet mastered, leading teachers to adopt whole-class re-teaching strategies.

Treatment teachers reported that time spent on developing a re-teaching plan, discussing instructional strategies or resources, and sharing ideas or resources during their data meetings was "somewhat" to "very" useful (not shown).

Leader Reported. Leaders in treatment and control schools both reported that a similar proportion of their teachers: 1) provided instruction to meet individual student needs, and 2) retaught students who were unsuccessful at mastering material the first time 3) using a different instructional approach (Table 17).

Compared to their counterparts in control schools, treatment school leaders more frequently reported that meetings to discuss data with teachers included time allocated to lesson planning ($d = 1.11, p = 0.00$) and sharing of instructional strategies ($d = 0.60, p = 0.01$; Table 17). Again, this is consistent with site visit data in which school leaders (and teachers) remarked that the additional focus on the steps that come *after* data analysis – i.e., the reteaching component – was a unique component of the ANet model compared to their past practices.

[Table 17]

3.4. Student Outputs & Additional Findings

In the first year, site visit and survey data collection focused primarily on the successes and difficulties associated with the implementation of ANet, including early indications of an impact of ANet on school leader and teacher practices. Fewer items were included to measure student outcomes such as improved academic motivation and academic achievement. This is not because student outcomes are less important. In fact, they are a key indicator in determining the impact of the ANet program. However, the analysis of student summative assessment scores will provide a better source of evidence of the ANet impact on student outcomes. In this section, we report preliminary survey-based findings related to student outcomes along with selected additional findings that do not align with specific components of the intervention logic model.

3.4.1. *Student Outputs & Outcomes*

Leader Reported. Fifty-three percent of leaders in control schools reported that students were able to review interim assessment results. In a common item, treatment school leaders were more likely to report that students received interim assessment results; however the difference was not statistically significant ($d = 0.33, p = 0.13$).

Teacher Reported. Sixty-seven percent of ANet teachers said they had seen changes in their students that they attribute to ANet. Preliminary examination of the write-in responses linked to this survey item indicates that the majority of changes were positive. For example, it was commonly noted that the ANet assessment data encouraged and motivated students to improve from one testing cycle to the next. They took greater ownership of their achievement and meeting achievement goals. However, a substantial minority of teachers' responses refer to the fact that the program increased students' exposure to testing, with adverse consequences for some students. As one teacher explained, "Some (students) were challenged and overwhelmed. Others welcomed the challenge and liked to compare scores."

These results support findings from year-one site visits. When asked about the impact of ANet on their students, a common teacher response was that they were more likely to share results with their students. The result of this sharing was that students had a greater awareness of their progress on particular standards, as well as how they compare to their classmates. One Boston teacher succinctly explained that sharing data

"... helps them take responsibility for their own learning, which I've never seen before. If they know, specifically, areas they need continued growth, they know where they can focus their energies to that and they know what they've mastered and what they are fine with. And it makes them want to work harder."

3.4.2. *Beliefs about Interim Assessments*

Teachers. Compared to the counterparts in control schools, treatment school teachers reported significantly higher agreement with statements about interim assessments keeping them accountable to their peers ($d = 0.25, p = 0.03$) and school leaders ($d = 0.28, p = 0.02$) for student progress (Table 18). This may be a reflection of the greater collaboration and leader follow-up experiences reported by treatment school teachers.

Treatment and control teachers responded similarly to items asking whether interim assessments place pressure on them to raise achievement and take time away from instruction. These results are reassuring given the criticism that an increased focus on testing can put pressure on teachers and result in the narrowing of the curriculum (Smith, 1991), and concerns raised by some teachers in ANet schools during site visit interviews:

“Sometimes the pace [of the assessment cycle] is difficult, to find the time to give it [the reteach] what it needs and deserves, because if you give it that time, then something else is compromised. Finding the balance, I guess, is the constant battle every teacher has.” (Springfield teacher)

The year-one data provides some evidence that the ANet treatment has not systematically led teachers to reduce their time on instruction. This is the case despite the fact that the administration of ANet interim assessments does appear to have taken up a substantial amount of class time. Although the ANet assessments are intended to be administered within an average class period (typically 40 to 50 minutes), site visit interviews indicated that it took students considerably more time to complete the assessments (especially noted in Jefferson Parish). In

fact, the mean treatment teacher response indicates that it took students about 1 hour 30 minutes to complete the math and ELA ANet interim assessments.

In a series of items about factors that hinder the use of interim assessment to plan or modify instruction, treatment school teachers were less likely to report that it was because of the absence of comparison data (Pr = -0.10, $p = 0.00$), because data are received too late (Pr = -0.15, $p = 0.00$) or at the wrong time of year to be useful (Pr = -0.06, $p = 0.10$)¹¹, or that it was because policies do not allow them instructional flexibility (Pr = -0.19, $p = 0.00$) (Table 18). However, treatment school teachers were more likely to report that it was because test data were not aligned with the state standards (Pr = 0.07, $p = 0.03$) or curriculum (Pr = 0.16, $p = 0.00$). The latter is consistent with responses to other items asking about the alignment of ANet interim assessments, though, for all teachers, these questions were asked generally of all interims. It is not a comparison between ANet and other interim assessments.

[Table 18]

3.4.3. Beliefs about Summative Assessments

School leaders and teachers were both asked about their agreement with a statement that end-of-year state tests only evaluate what is easy to measure, not what is important. Within each group, there was no treatment/control difference in response patterns (Table 18). However, the mean responses indicate higher agreement from teachers, perhaps reflecting greater discomfort among teachers with the reliance on standardized tests to measure student learning.

3.4.4. School Culture

¹¹ $p = .111$ for the OLS model, $p = .087$ for the ordered probit model

School leaders. Leaders in both groups responded similarly to a series of questions about the proportion of their teaching staff that: acted in a way that communicates an investment in improving their teaching, collaborates with other teachers around instruction, acts in a manner that communicates all students can learn, and sets high expectations for students. Average responses indicate that between “about half” and “many” teachers fit these characteristics.

Teachers. A set of items asking about the proportion of the teaching staff that held certain beliefs or acted in certain ways: feeling responsible to help each other or when students fail, really try to improve their teaching, take risks to make the school better, and motivate students to learn. Teachers in both treatment and control schools responded similarly, with the mean response generally falling between “some” and “most” (Table 19).

[Table 19]

4. Conclusion

This report, the first in CEPR's ongoing evaluation of the Achievement Network's i3 expansion, uses unusually rich survey data to provide systematic evidence on how ANet influences the practice of school leaders and teachers during their first year of program participation. In particular, by comparing the responses of educators in schools randomly assigned to ANet participation to those of educators in a control group of non-participating schools, we shed light on how the ANet model differs from standard approaches to instructional data use in the types of urban schools in which the organization is currently expanding. Understanding these differences is critical, as our broader evaluation of ANet's impact on student achievement will isolate the program's impact over and above these standard approaches.

Our analysis identifies several differences that appear to reflect key priorities of the ANet model. Perhaps most notably, both leader and teacher responses indicate that the interim assessments provided by ANet are considerably more rigorous than the interim assessments administered in control schools. Leaders also report that the ANet interim assessments are better aligned with both the state standards and the curricular scope and sequence in use in their district or school. Interestingly, however, teachers report that the ANet interim assessments are no better aligned with the state standards and that they are less aligned to their curricular scope and sequence in mathematics.

Consistent with ANet's heavy investment in training and logistical support, both leaders and teachers rate the supports available for instructional data use far more highly than do leaders and teachers in control schools. Leaders in ANet schools also report greater confidence in their ability to use interim assessment data for several purposes: setting challenging goals for school-wide achievement; comparing their school's performance with that of other schools serving

similar students; examining trends in their school's performance over time; and identifying students in need of instructional support. In contrast, leader confidence in their ability to use interim assessment data was no higher in two areas not explicitly incorporated into ANet protocols: evaluating the performance of individual teachers or the effectiveness of instructional programs.

We also find suggestive evidence that ANet participation causes school leaders to take steps to ensure that interim assessment data influences teachers' classroom practices. For example, ANet school leaders report that their staff members responsible for reviewing interim assessment data were more likely to observe teacher practice, meet with teachers to discuss interim assessment data, coach teachers in using interim assessment data to monitor student progress, and coach teachers in developing re-teaching plans. Teachers in ANet schools also report that their teachers are better able to monitor the quality of teaching, institute concrete practices and procedures to encourage data use, and provide actionable feedback on lesson plans.

Finally, teachers in ANet schools report that, during common planning time, they more often use data in certain ways emphasized by the ANet protocol. Specifically, they report that they use data in common planning time more frequently to identify when and what to re-teach, determine whether a lesson should be delivered to the entire class or to a subset of students, and to identify supplemental curricular materials. Teachers in ANet schools also report that they more frequently use interim assessment data on their own to plan their instruction and measure student progress toward identified goals. Conversely, we observe no differences between teachers in ANet and control schools in their use of data in common planning time to identify which students need instructional support or in their personal use of interim assessment data to adjust their teaching plans to better meet students' learning needs.

On these and other items where we find no evidence that ANet has impacted leader or teacher behavior, it is important to keep in mind that this does not necessarily imply that educators in ANet schools are not engaging in these practices. Rather, it may simply be the case that educators in control schools also do so. It is also possible that participating in the ANet program could alter *how* educators engage in a given practice in ways that are not captured by our survey. Finally, it is important to emphasize that our survey was administered during the first year of the treatment schools' participation with ANet. It is expected that program impacts on educator behavior may change as schools move deeper into their partnership. A second survey administered in the Spring of the 2012-13 school year will allow us to explore this possibility.

5. References

- Carlson, D., Borman, G. D., & Robinson, M. (2011). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis*, 33(3), 378-398.
- Clune, W. H., & White, P. A. (2008). *Policy effectiveness of interim assessments in Providence public schools* (WCER Working Paper No. 2008-10). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research.
- Dobbie, W. & Fryer, Jr., R. G. (2011). *Getting beneath the veil of effective schools: Evidence from New York City* (Working Paper No. 17632). Cambridge: National Bureau of Economic Research. .
- Henderson, S., Petrosino, A., Guckenburg, S., & Hamilton, S. (2007). *Measuring how benchmark assessments affect student achievement*, (Issues & Answers Report, REL 2007 No. 039). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- Henderson, S., Petrosino, A., Guckenburg, S., & Hamilton, S. (2008). *REL Technical Brief—a second follow-up year for ‘Measuring how benchmark assessments affect student achievement,’* (REL Technical Brief, REL Northeast and Islands 2008–No. 002). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- Imai, K., King, G., & Nall, C. (2009). The essential role of pair matching in cluster-randomized experiments, with application to the Mexican universal health insurance evaluation. *Statistical Science*, 24(1): 29-53.
- Kerr, K. K., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons from three urban districts. *American Journal of Education*, 112(5): 496-520.
- Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues and Practice*, 28(3), 5–13.
- Quint, J. C., Sepanik, S., & Smith, J. K. (2008) *Using student data to improve teaching and learning: Findings from an evaluation of the Formative Assessments of Student Thinking in Reading (FAST-R) in Boston elementary schools*. New York: MDRC.
- Riggan, M. & Olah, L. N. (2011). Locating interim assessments within teachers’ assessment practice, *Educational Assessment*, 16(1):1–14.

Shepard, L. A. (2010): What the marketplace has brought us: Item-by-item teaching with little instructional insight. *Peabody Journal of Education*, 85(2): 246-257.

Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11.

Stecher, B. M., Epstein, S., Hamilton, L. S., Marsh, J. A., Robyn, A., McCombs, J. S., Russell, J., & Naftel, S. (2008). *Pain and gain: Implementing No Child Left Behind in three states, 2004–2006*. Santa Monica, CA: RAND Corporation.

Appendix A. Wave One Samples

For the first wave of the Achievement Network (ANet) i3 study, schools were recruited into the sample from Boston Public Schools (MA), Chelsea Public Schools (MA), Chicago Public Schools (IL), Jefferson Parish Public Schools (LA), and Springfield Public Schools (MA). These school districts were chosen based on prior plans to expand the total number of schools ANet served in geographically concentrated networks. ANet worked with these districts to establish a memorandum of understanding (MOU) that they would provide subsidized services to a specific number of schools in each district. Individual schools were recruited based on their interest in working with ANet and their responses to a screening application. The resulting recruitment effort netted 101 schools (treatment $n = 51$; control school $n = 50$). Schools were paired on a series of school-level demographic and performance characteristics and randomly assigned to the treatment or control group. With an odd number of schools in Chelsea, one “pair” consists of three schools; two assigned to the treatment and one assigned to the control group.

School-Level Exclusions

The full sample of 101 schools was reduced due the withdrawal and reorganization of schools in several districts and resulting in a smaller primary impact sample (i.e., student impact sample). After recruitment had been completed in three regional Areas within Chicago Public Schools and 36 schools had been randomized into treatment or control conditions (but prior to any implementation of the intervention), the district experienced a leadership transition and an internal reorganization. Because the district’s schools lost control over the discretionary budgets they had planned to use to pay for ANet services, ANet had to negotiate a revised MOU with Chicago Public Schools. This revised MOU enabled ANet to administer its program of 10 of the

original 18 schools that had been randomly assigned to the treatment condition. Eight schools and their matched pairs were dropped from the primary impact sample.

After randomization was completed in Boston Public Schools (but prior to any implementation of services), two BPS schools withdrew from the study. They and their pairs were dropped from the primary impact sample. In addition, ANet decided that maintaining its partnership with the district required that it serve two schools in years 1 and 2 that had initially been assigned to the control condition. As a result, we dropped these two schools and their matched pairs from the primary impact sample, but are continuing to gather data from the schools so that they can be included in exploratory analysis probing the robustness of our impact results to this deviation from the planned assignment process. Finally, one BPS control school refused to participate in the survey data collection. They and their pair were dropped from the *teacher-leader* impact sample only. These two schools are still included in the primary impact sample (i.e., student outcomes impact sample).

One school in JPPSS (and its pair) was dropped from the primary impact sample, because it is an alternative school serving a unique population of at-risk students. One school in SPS withdrew from the study prior to implementation. This school and its pair are excluded in the primary impact sample. In total, 28 schools were removed from the primary impact sample and 30 from the teacher-leader impact sample; thus, resulting in the 71 schools – 35 control schools and 36 treatment schools – used as the basis for the sample in this report.

Table A1. School-Level Exclusions

District	Sample(s) Excluded From ¹	Number Excluded
CPS		
Eight schools and their pairs dropped due to district reorganization and loss of discretionary funding	PI and TL	16
BPS		
Two schools withdrew from the study; they and their pairs were removed	PI and TL	4
Two control schools received the treatment; they and their pairs were removed	PI and TL	4
One school refused to participate in survey data collection; they and their pair were removed	TL only	2
JPPSS		
One school was dropped due to grade configuration; they and their pair were removed	PI and TL	2
SPS		
One school withdrew from the study; they and their pair were removed	PI and TL	2
Total Teacher-Leader Impact Sample Exclusions		30

¹ Samples include the primary impact sample (PI) and teacher-leader impact sample (TL).

Individual-Level Exclusions & the Year-One Analytic Samples

School leaders. A total of 64 year-one surveys were received from schools leaders; however, 7 surveys were received from leaders in schools that had withdrawn or been otherwise excluded from the teacher-leader impact sample. No other exclusions were enforced on the leader analytic sample. Of the 71 schools in the teacher-leader impact sample, the leader analytic sample consisted of responses from 57 school leaders representing 57 schools (i.e., one leader response per school). The breakdown of responses was evenly split with responses from 29 control schools and 28 treatment schools.

Teachers. A total of 755 year-one surveys were received from teachers; however, 101 surveys were received from leaders in schools that had withdrawn or been otherwise excluded from the teacher-leader analytic sample. The remaining 654 teachers represented all 71 schools in the teacher-leader impact sample: 236 responses from control-school teachers and 418 responses from treatment-group teachers. However, the teacher analytic sample was subject to two additional exclusions: teachers who changed schools ($n = 2$) and teachers who were “out-of-scope” given the study’s population of interest as discussed in the text.

Table 1. Baseline Equivalence: Tests of Significant Differences Between Treatment and Control Schools

	Treatment Group	Control Group	Difference in	OLS Coeff.	p-value
	Mean	Mean	Means		
SCHOOL DEMOGRAPHICS					
Percent Free and Reduced Price Lunch Enrollment	88.2 (8.32)	88.1 (9.96)	0.2 (2.18)	0.29 (1.87)	0.88
Percent English Language Learner Enrollment	32.9 (31.78)	31.1 (32.87)	1.9 (9.64)	-0.401 (4.66)	0.93
Percent Special Education Enrollment	14.3 (6.43)	14.4 (6.51)	0.0 (1.54)	-0.06 (1.17)	0.96
Percent White Enrollment	25.0 (31.10)	25.6 (30.68)	-0.6 (7.33)	-0.14 (6.64)	0.98
Percent Black Enrollment	39.1 (34.25)	39.2 (34.82)	-0.1 (8.20)	0.79 (7.73)	0.92
Percent Hispanic Enrollment	31.7 (29.42)	31.5 (28.94)	0.2 (6.93)	-1.23 (3.64)	0.74
Percent Other Race/Ethnicity Enrollment	3.8 (5.00)	3.0 (3.37)	0.7 (1.02)	0.75 (0.85)	0.38
School Size	481.1 (192.24)	422.1 (145.29)	59.0 (40.53)	57.12 (40.40)	0.16
Total	36	35	71	71	
SCHOOL PERFORMANCE					
Grade 3 Math	41.0 (22.87)	40.7 (19.93)	0.3 (5.28)	0.3 (4.02)	0.94
Total	33	33	66	66	
Grade 3 Reading	34.3 (18.49)	34.7 (17.93)	-0.3 (4.48)	-0.3 (3.55)	0.92
Total	33	33	66	66	
Grade 6 Math	42.3 (17.17)	48.5 (15.73)	-6.2 (6.13)	-5.8 (5.77)	0.32
Total	15	14	29	29	
Grade 6 Reading	46.9 (13.31)	55.0 (10.84)	-8.1 (4.53)	-8.0 (4.24)	0.07
Total	15	14	29	29	

Note: Total sample = 71 schools.

ELL totals differ from the totals reported since not all schools had ELL information (treatment group n=23, control group n=22).

Treatment and control group means and difference in means are based on a two-group mean-comparison test.

The OLS coefficient and p-value is the result of the regression of the demographic/performance variable on treatment status with statistical controls for network and whether they belong to the pair with three schools.

Numbers in parentheses are standard deviations (for the treatment group mean column and control group mean column) and standard errors (for the difference in means column and OLS coeff column).

Source: School demographic data from 2008-09 Common Core of Data, school performance data from 2009-10 Common Core of Data.

Table 2. Descriptives on Sample Respondents (Leaders)

	Control Group		Treatment Group	
	Estimate	N	Estimate	N
LEADERS				
Gender (percentage)				
Female	82.8	24	78.6	22
Male	10.3	3	17.9	5
Missing	6.9	2	3.6	1
Race/Ethnicity (percentage)				
American Indian or Alaskan Native	--	--	--	--
Asian or Pacific Islander	--	--	--	--
Black or African American	48.3	14	35.7	10
Hispanic or Latino	3.5	1	7.1	2
White	41.4	12	53.6	15
Other	--	--	--	--
Missing	6.9	2	3.6	1
Highest Level of Education (percentage)				
Master's degree	89.7	26	89.3	25
Doctorate	3.5	1	10.7	3
Missing	6.9	2	--	--
Years of Experience (mean)				
Total (leadership)	5.7		6.1	
(SE)	(1.15)	27	(0.69)	28
This district (total)	18.3		19.7	
(SE)	(1.96)	27	(1.85)	28
This school (leadership)	3.6		4.8	
(SE)	(0.60)	27	(0.61)	28

Note: Total leader n = 57. Leaders were asked to report number of years as principal total and in their current school, but total number of years they worked in the district. This could include time as a teacher, assistant principal, etc.

Source: Year-one (2011-12) school leader survey.

Table 3. Descriptives on Sample Respondents (Teachers)

	Control Group		Treatment Group	
	Estimate	N	Estimate	N
TEACHERS				
Gender (percentage)				
Female	88.2	165	88.5	301
Male	11.2	21	9.7	33
Missing	0.5	1	1.8	6
Race/Ethnicity (percentage)				
American Indian or Alaskan Native	--	--	0.3	1
Asian or Pacific Islander	1.6	3	2.4	8
Black or African American	15.5	29	24.1	82
Hispanic or Latino	7.0	13	4.4	15
White	70.6	132	64.1	218
Other	3.7	7	2.7	9
Missing	1.6	3	2.1	7
Highest Level of Education (percentage)				
Bachelor's degree	31.6	59	32.7	111
Master's degree	67.9	127	64.4	219
Doctorate	0.5	1	1.5	5
Missing	--	--	1.5	5
Alternative Certification (percentage)				
No	82.4	154	78.5	267
Yes	17.1	34	20.6	70
Missing	0.5	1	0.9	3
Grade Levels Taught (percentage)				
3rd	38.5	72	33.2	113
4th	36.9	69	32.7	111
5th	33.2	62	32.4	110
6th	13.9	26	17.9	61
7th	15.5	29	21.2	72
8th	14.4	27	17.9	61
Other	4.8	9	4.1	14
Subject Areas Taught (percentage)				
English language arts or reading	45.5	85	43.8	149
Mathematics	34.2	64	39.1	133
Elementary/general education	27.8	53	29.7	101
Special education	17.1	32	15.0	51
English as a second language/English language development	9.6	18	9.1	31
Other	11.8	22	7.1	24
Years of Teaching Experience (mean)				
Total	12.7		11.7	
(SE)	(0.64)	186	(0.48)	336
This district	10.1		10.4	
(SE)	(0.53)	186	(0.47)	337
This school	7.0		6.8	
(SE)	(0.42)	185	(0.36)	337
Current grade or subject	6.6		6.3	
(SE)	(0.44)	184	(0.33)	336

Note: Total teacher n = 527.

Grade level and subject area taught were 'select all' items. Teachers could report multiple grades or subjects. Total will sum to greater than 100 percent.

Source: Year-one (2011-12) teacher survey.

Table 4. Teacher Attrition: Tests of Significant Differences Between Treatment and Control Schools in Teacher Analytic Sample

	Treatment Group Mean	Control Group Mean	Difference in Means	OLS Coeff.	p -value
SCHOOL DEMOGRAPHICS					
Percent Free and Reduced Price Lunch Enrollment	88.2 (8.32)	87.3 (10.46)	0.9 (2.31)	0.4 (2.00)	0.83
Percent English Language Learner Enrollment	32.9 (31.78)	35.6 (33.18)	-2.6 (10.05)	-0.3 (5.04)	0.95
Percent Special Education Enrollment	14.3 (6.43)	15.0 (6.60)	-0.7 (1.61)	-0.4 (1.25)	0.73
Percent White Enrollment	25.0 (31.10)	22.3 (27.84)	2.8 (7.33)	1.6 (6.74)	0.82
Percent Black Enrollment	39.1 (34.25)	40.4 (34.44)	-1.3 (8.49)	-1.3 (7.89)	0.88
Percent Hispanic Enrollment	31.7 (29.42)	33.6 (29.86)	-1.9 (7.32)	-1.1 (3.90)	0.77
Percent Other Race/Ethnicity Enrollment	3.8 (5.00)	3.2 (3.43)	0.6 (1.08)	1.0 (0.92)	0.30
School Size	481.1 (192.24)	413.6 (143.71)	67.5 (42.51)	63.0 (43.02)	0.15
Total	36	30	66	66	
SCHOOL PERFORMANCE					
Grade 3 Math	41.0 (22.87)	38.3 (18.26)	2.7 (5.37)	1.5 (4.24)	0.73
Total	33	28	61	61	
Grade 3 Reading	34.3 (18.49)	33.3 (17.43)	1.0 (4.63)	0.2 (3.73)	0.96
Total	33	28	61	61	
Grade 6 Math	42.3 (17.17)	46.4 (16.51)	-4.2 (6.71)	-4.9 (6.40)	0.45
Total	15	11	26	26	
Grade 6 Reading	46.9 (13.31)	51.9 (9.88)	-5.1 (4.76)	-5.8 (4.56)	0.22
Total	15	11	26	26	

Note: Total sample = 66 schools.

ELL totals differ from the totals reported since not all schools had ELL information (treatment group n=23, control group n=19).

Treatment and control group means and difference in means are based on a two-group mean-comparison test.

The OLS coefficient and p-value is the result of the regression of the demographic/performance variable on treatment status with statistical controls for network and whether they belong to the pair with three schools.

Numbers in parentheses are standard deviations (for the treatment group mean column and control group mean column) and standard errors (for the difference in means column and OLS coeff column).

Source: School demographic data from 2008-09 Common Core of Data, school performance data from 2009-10 Common Core of Data.

Table 5. Leader Attrition: Tests of Significant Differences Between Treatment and Control Schools in School Leader Analytic Sample

	Treatment Group Mean	Control Group Mean	Difference in Means	OLS Coeff.	p-value
SCHOOL DEMOGRAPHICS					
Percent Free and Reduced Price Lunch Enrollment	87.9 (8.37)	87.6 (10.00)	0.4 (2.45)	0.4 (2.18)	0.87
Percent English Language Learner Enrollment	41.2 (32.32)	39.3 (33.14)	1.9 (11.23)	0.1 (6.02)	0.98
Percent Special Education Enrollment	14.5 (6.73)	15.2 (6.60)	-0.7 (1.77)	-0.3 (1.36)	0.81
Percent White Enrollment	24.4 (30.30)	16.6 (20.27)	7.8 (6.81)	7.7 (6.33)	0.23
Percent Black Enrollment	34.9 (30.37)	43.0 (33.22)	-8.2 (8.44)	-8.2 (7.40)	0.27
Percent Hispanic Enrollment	35.9 (30.01)	36.1 (29.40)	-0.2 (7.87)	-0.3 (3.98)	0.93
Percent Other Race/Ethnicity Enrollment	4.6 (5.34)	3.6 (3.44)	1.0 (1.19)	1.1 (1.05)	0.30
School Size	453.5 (155.42)	391.0 (132.62)	62.5 (38.22)	57.7 (39.13)	0.15
Total	28	29	57	57	
SCHOOL PERFORMANCE					
Grade 3 Math	37.2 (20.24)	37.9 (19.16)	-0.7 (5.46)	-0.2 (4.38)	0.96
Total	25	27	52	52	
Grade 3 Reading	30.2 (13.94)	32.7 (18.39)	-2.5 (4.55)	-1.9 (3.64)	0.61
Total	25	27	52	52	
Grade 6 Math	37.5 (13.46)	45.4 (17.06)	-8.0 (7.24)	-8.3 (7.66)	0.30
Total	9	9	18	18	
Grade 6 Reading	44.3 (8.98)	51.0 (10.90)	-6.7 (4.71)	-7.4 (4.93)	0.16
Total	9	9	18	18	

Note: Total sample = 57 schools.

ELL totals differ from the totals reported since not all schools had ELL information (treatment group n=17, control group n=17).

Treatment and control group means and difference in means are based on a Two-group mean-comparison test.

The OLS coefficient and p-value is the result of the regression of the demographic/performance variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Numbers in parentheses are standard deviations (for the treatment group mean column and control group mean column) and standard errors (for the difference in means column and OLS coeff column).

Source: School demographic data from 2008-09 Common Core of Data, school performance data from 2009-10 Common Core of Data.

Exhibit 2. Organization of Findings, by Logic Model Component

	Not Measured	Source Data			
		School Leader		Teacher	
		Treatment	Control	Treatment	Control
Intervention Inputs					
Aligned assessment		✓	✓	✓	✓
Logistical support		✓	✓	✓	✓
Training and coaching		✓	✓	✓	✓
Network activities		✓	✓	--	--
Leadership Actions					
Establish data leadership team	✓	--	--	--	--
Establish goals for student achievement		✓	✓	✓	✓
Manage assessment administration	✓	--	--	--	--
Analyze assessment results		✓	✓	✓	✓
Facilitate data meetings		✓	--	--	--
Ensure teachers take instructional action and reflect		✓	✓	✓	✓
Use data to build positive school culture	✓ ¹				
Teacher Actions					
Establish goals for student achievement		--	--	✓	✓
Analyze assessment results		✓	✓	✓	✓
Use backward planning, develop and implement action plans, and monitor and reflect on action plans ²		✓	✓	✓	✓
Student Outputs and Outcomes					
		✓	✓	✓	--

¹ Many of the items in other sections of the findings address issues around school culture; however, few items were specific enough to fall under leadership actions: use data to build positive school culture. For more general school culture items, see section 3.4.4. of the report findings.

² These are separate logic model components, but discussed in the findings in a single section (3.3.3.).

Table 6. Aligned Assessments: Alignment and Rigor of Interim Assessments

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i> -value	N
LEADERS					
Thinking about the ELA interim assessments you administered this year, please rate their alignment with:					
Your state's ELA standards	3.65	0.44 **	0.70	0.02	54
Your state's end-of-year ELA assessment	3.54	0.52 ‡	0.89	0.01	54
Your school or district's ELA curriculum	3.37	0.35 *	0.47	0.09	55
Your school or district's curricular scope and sequence for ELA	3.30	0.44 **	0.57	0.04	55
Thinking about the ELA interim assessments you administered this year, please rate the level of rigor of test items compared to:					
Your state's ELA standards	2.74	0.95 ‡	1.33	0.00	55
Your state's end-of-year ELA assessment	2.67	1.17 ‡	1.40	0.00	54
Your school or district's ELA curriculum	2.89	0.99 ‡	1.24	0.00	55
Thinking about the mathematics interim assessments you administered this year, please rate their alignment with:					
Your state's mathematics standards	3.43	0.67 ‡	1.16	0.00	54
Your state's end-of-year mathematics assessment	3.48	0.71 ‡	1.22	0.00	53
Your school or district's mathematics curriculum	3.44	0.37 *	0.57	0.07	53
Your school or district's curricular scope and sequence for mathematics	3.46	0.27	0.42	0.22	54
Thinking about the mathematics interim assessments you administered this year, please rate the level of rigor of test items compared to:					
Your state's mathematics standards	2.86	0.87 ‡	1.48	0.00	55
Your state's end-of-year mathematics assessment	2.75	1.00 ‡	1.70	0.00	53
Your school or district's mathematics curriculum	2.86	1.05 ‡	2.00	0.00	55
TEACHERS					
Thinking about the ELA interim assessments you administered this year, please rate their alignment with:					
Your state's ELA standards	3.87	0.03	0.05	0.72	439
Your state's end-of-year ELA assessment	3.79	-0.01	-0.01	0.94	438
Your school or district's ELA curriculum	3.81	-0.14	-0.18	0.15	438
Your school or district's curricular scope and sequence for ELA	3.72	-0.12	-0.15	0.31	436
Thinking about the ELA interim assessments you administered this year, please rate the level of rigor of test items compared to:					
Your state's ELA standards	3.06	0.40 ‡	0.57	0.00	434
Your state's end-of-year ELA assessment	3.01	0.48 ‡	0.63	0.00	433
Your school or district's ELA curriculum	3.07	0.55 ‡	0.76	0.00	431
Thinking about the mathematics interim assessments you administered this year, please rate their alignment with:					
Your state's mathematics standards	3.86	0.00	0.00	0.98	377
Your state's end-of-year mathematics assessment	3.81	-0.01	-0.01	0.92	378
Your school or district's mathematics curriculum	3.76	-0.18 *	-0.23	0.09	377
Your school or district's curricular scope and sequence for mathematics	3.79	-0.35 ‡	-0.43	0.00	375
Thinking about the mathematics interim assessments you administered this year, please rate the level of rigor of test items compared to:					
Your state's mathematics standards	3.15	0.33 ‡	0.55	0.00	379
Your state's end-of-year mathematics assessment	3.17	0.30 ‡	0.43	0.00	375
Your school or district's mathematics curriculum	3.15	0.44 ‡	0.69	0.00	376

Note: Total leader n = 57; total teacher n = 527. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Standard errors were clustered by school for all teacher-level analyses. In responding to all items, treatment-school leaders and teachers were asked to consider the ANet interim assessments they administered. Control-school leaders and teachers were asked to consider all of the interim assessments they administered.

Alignment items response scale: 1=very poor, 2=poor, 3=fair, 4=good, 5=excellent. Rigor items response scale: 1=much less rigorous, 2=less rigorous, 3=about the same level of rigor, 4=more rigorous, 5=much more rigorous.

* significant at .10; ** significant at .05; ‡ significant at .01.

Source: Year-one (2011-12) school leader and teacher surveys.

Table 7. Logistical Support: MyANet (Treatment only)

	Treatment		
	Group Mean	St. Dev.	N
LEADERS			
Are you aware of ANet's online tools, "MyANet?"	0.96	0.19	28
How often have you used the following online tools on MyANet?			
Quiz generator	2.04	1.09	27
Better lessons	2.08	1.13	26
Tools for reviewing students' scores on interim assessments	3.19	0.90	26
Tools for comparing students' scores across classes at this schools	3.08	0.74	26
Tools for comparing students' scores with other schools in my network	2.88	1.01	25
List of state ELA or mathematics standards	2.46	1.07	26
Interim assessment materials	2.27	0.92	26
How useful do you find the following online tools on MyANet?			
Quiz generator	2.67	0.49	15
Better lessons	2.79	0.43	14
Tools for reviewing students' scores on interim assessments	2.92	0.28	24
Tools for comparing students' scores across classes at this schools	2.92	0.28	24
Tools for comparing students' scores with other schools in my network	2.82	0.39	22
List of state ELA or mathematics standards	2.67	0.49	18
Interim assessment materials	2.72	0.46	18
TEACHERS			
Are you aware of ANet's online tools, "MyANet?"	0.97	0.17	340
How often have you used the following online tools on MyANet?			
Quiz generator	2.90	1.01	328
Better lessons	2.42	1.16	328
Tools for reviewing students' scores on interim assessments	3.05	0.89	329
Tools for comparing students' scores across classes at this schools	2.80	1.00	329
Tools for comparing students' scores with other schools in my network	2.56	1.06	329
List of state ELA or mathematics standards	2.82	1.06	329
Interim assessment materials	2.69	1.03	323
How useful do you find the following online tools on MyANet?			
Quiz generator	2.69	0.48	293
Better lessons	2.41	0.57	231
Tools for reviewing students' scores on interim assessments	2.67	0.47	305
Tools for comparing students' scores across classes at this schools	2.56	0.52	282
Tools for comparing students' scores with other schools in my network	2.49	0.58	255
List of state ELA or mathematics standards	2.54	0.55	279
Interim assessment materials	2.49	0.54	270

Note: Total treatment leader n = 29; total treatment teacher n = 340. Use items response scale: 1=never, 2=once or twice a year, 3=monthly, 4=weekly, 5=more than once a week. Usefulness items response scale: 1=not at all useful, 2=somewhat useful, 3=very useful.

Source: Year-one (2011-12) school leader and teacher surveys, treatment items only.

Table 8. Training and Support: Level and Satisfaction with Support Received

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i> -value	N
LEADERS					
Please rate the level of support you received in the use of interim assessment data from:					
Your ANet coach (treatment)/an external provider (control)	2.00	0.93 ‡	1.00	0.00	52
You school or district	2.41	-0.09	-0.12	0.64	55
How satisfied are you with the overall quality of the support you receive in the following?¹					
Analysis of assessment data	3.26	1.19 ‡	1.26	0.00	55
Setting student performance goals	3.30	0.86 ‡	0.94	0.00	55
Timeliness of response to questions	3.48	1.15 ‡	1.29	0.00	55
Improving instructional practices to meet students' needs	3.00	1.16 ‡	1.11	0.00	54
TEACHERS					
How satisfied are you with the resources you have at your current school to:					
Meet students' individual needs	3.29	0.16	0.16	0.26	523
Track individual student progress	3.62	0.15	0.16	0.25	523
Please rate the level of support you received in the use of interim assessment data from:					
Your ANet coach (treatment)/an external provider (control)	2.54	0.20 **	0.26	0.04	508
You school or district	2.11	0.54 ‡	0.62	0.00	503
How satisfied are you with the overall quality of the support you receive in the following?¹					
Analysis of assessment data	3.43	0.49 ‡	0.47	0.00	510
Setting student performance goals	3.42	0.39 ‡	0.39	0.01	509
Timeliness of response to questions	3.39	0.51 ‡	0.51	0.00	505
Improving instructional practices to meet students' needs	3.32	0.38 **	0.36	0.01	504

Note: Total leader n = 57; total teacher n = 527. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Standard errors were clustered by school for all teacher-level analyses.

¹ In responding to these items, treatment-school leaders and teachers were asked to consider their ANet coach. Control-school leaders and teachers were more generally asked about the quality of support they received

Satisfaction items response scale: 1=very dissatisfied 2=dissatisfied, 3=neither satisfied nor dissatisfied 4=satisfied, 5=very satisfied. Level of support items scale: 1=not nearly enough, 2=would have liked a little more, 3=amount of support was adequate, 4=more support than I needed.

** significant at .05; ‡ significant at .01.

Source: Year-one (2011-12) school leader and teacher surveys.

Table 9. Professional Development (Teachers Only)

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i> -value	N
TEACHERS					
Thinking about all of your professional development experiences this year, please indicate the extent to which you agree or disagree with the following.					
Overall, my professional development experiences have:					
Been sustained and coherently focused, rather than short-term and unrelated.	3.51	0.12	0.11	0.30	527
Included enough time to think carefully about, try, and evaluate new ideas.	3.40	0.12	0.11	0.36	526
Been closely connected to my school's improvement plan.	3.75	0.03	0.03	0.84	525
Helped me build new skills to better meet my students' learning needs.	3.62	0.18 *	0.18	0.07	525
Included opportunities to work productively with colleagues in my school.	3.49	0.29 **	0.26	0.02	526
Included opportunities to work productively with teachers from other schools.	2.73	0.12	0.09	0.35	525
Included follow-up support from leaders to implement what was learned.	2.99	0.32 *	0.28	0.05	519

Note: Total teacher n = 527. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Standard errors were clustered by school for all teacher-level analyses.

Agreement items response scale: 1=strongly disagree, 2=disagree, 3=neither agree nor disagree, 4=agree, 5=strongly agree.

* significant at .10; ** significant at .05.

Source: Year-one (2011-12) teacher survey.

Table 10. Leader Actions: Establish Goals (Leader-reported Only)

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	Marginal Effect	<i>p</i> -value	N
LEADERS						
Goal setting						
Earliest occurrence of goal setting	2.04	-0.07	-0.05	--	0.86	52
Number of time goals were set during school year	2.04	-0.53	-0.29	--	0.20	52
At what level did you set goals for student performance on the state test						
Schoolwide	0.85	-0.04	--	-0.06	0.71	52
Grade level	0.85	-0.12	--	-0.12	0.30	52
Classroom level	0.58	-0.13	--	-0.14	0.37	52
Individual students	0.50	-0.08	--	-0.11	0.52	52
What factors helped you set goals for student performance?						
Statewide targets to meet AYP	0.81	0.11	--	0.14	0.26	52
School-specific targets to meet AYP via Safe Harbor	0.54	0.00	--	0.00	0.99	52
District performance goals	0.65	0.08	--	0.09	0.52	52
Current student achievement levels	0.85	0.04	--	0.04	0.70	52
Comparable data from other schools	0.15	0.07	--	0.05	0.53	52
Interim assessment data	0.65	-0.04	--	-0.05	0.77	52
How confident are you in using interim assessment data to set challenging goals for student achievement school-wide						
	3.63	0.43 **	0.58	--	0.03	55

Note: Total leader n = 57. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools.

Earliest goal setting item response options: 1=prior to school year, 2=Sept, 3=Oct, 4=Nov, 5=Dec, 6=Jan, 7=Feb, 8=March, 9=April, 10=May.

Levels and factors items response options: 0=no, 1=yes.

** significant at .05.

Source: Year-one (2011-12) school leader survey.

Table 11. Leader Actions: Analyze Results (Leader-reported Only)

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	Marginal Effect	<i>p</i> -value	N
LEADERS						
Did your school use interim assessment data this year to:						
Evaluate programs	0.44	0.09	--	0.10	0.49	55
Evaluate teachers	0.37	-0.08	--	-0.11	0.49	55
Compare the performance of different grades and/or classrooms	0.70	0.10	--	0.14	0.38	54
Compare the performance of different groups of students	0.63	0.12	--	0.12	0.34	55
Compare the school's performance with other schools	0.52	0.41 ‡	--	0.42	0.00	55
Refer students for additional services	0.81	0.08	--	0.12	0.40	55
How confident are you in using interim assessment data to perform each of the activities below:						
Compare your school's performance to that of other schools	3.56	0.52 **	0.61	--	0.02	55
Examine trends in your school's performance over time	3.59	0.58 ‡	0.73	--	0.01	55
Evaluate the performance of individual teachers in raising student achievement	3.81	0.06	0.08	--	0.79	55
Identify struggling students in need of instructional support	3.70	0.38 *	0.48	--	0.06	55
Evaluate the effectiveness of instructional programs	4.07	-0.29	-0.37	--	0.16	55

Note: Total leader *n* = 57. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools.

Use item scale: 0=no, 1=yes. Confidence items scale: 1=not at all confident, 2=a little confident, 3=somewhat confident, 4=quite confident, 5=highly confident.

* significant at .10; ** significant at .05; ‡ significant at .01.

Source: Year-one (2011-12) school leader survey.

Table 12. Leader Actions: Facilitate Data Meetings (Treatment Leader-reported Only)

	Treatment Group Mean	St. Dev.	N
LEADERS			
Thinking about data meetings, please rate how useful it was to:			
Review your class' results from the ANet ELA or math interim assessment	2.89	0.315	28
Review individual student responses from the ANet ELA or math interim assessment	2.81	0.396	27
Identify a priority skill/standard	2.82	0.390	28
Develop a lesson plan to re-teach a specific skill/standard	2.75	0.441	28
Discuss instructional strategies or resources for re-teaching a specific skill/standard with teachers in my grade or content area	2.78	0.424	27
Share re-teaching ideas or resources with teachers from other grade levels or content areas	2.78	0.424	27

Note: Total leader n = 29. Items response scale: 1=not at all useful, 2=somewhat useful, 3=very useful.

Source: Year-one (2011-12) school leader survey, treatment items only.

Table 13. Leader Actions: Ensure Teacher Actions

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i> -value	N
LEADERS					
How often did you or another instructional leader engage in the following activities:					
Observing classroom instruction	4.36	0.33	0.32	0.17	56
Participating in grade-level or content area team meetings	4.82	-0.09	-0.22	0.48	56
Developing individual professional development plans with teachers	2.64	0.41	0.38	0.18	55
Implementing activities that encourage teachers to reflect on their instructional practice	3.39	0.66 **	0.55	0.03	56
Thinking about the person/people on your staff directly responsible for reviewing interim assessment data, on average, how often did they:					
Observe the instruction of individual teachers	3.59	0.54 *	0.43	0.09	55
Meet with teachers to discuss student assessment data	3.81	0.43 *	0.54	0.06	54
Meet with teachers to discuss students who are not meeting grade level expectations	3.73	0.06	0.06	0.85	53
Discuss teachers' unit and/or lesson plans	3.74	-0.09	-0.08	0.77	55
Coach teachers in planning backwards from standards	2.63	0.45	0.29	0.28	55
Coach teachers in using student assessment data to monitor student progress	3.04	0.73 **	0.55	0.02	55
Coach teachers in developing plans to re-teach content to struggling students	3.22	0.49 *	0.38	0.10	55
Coach teachers in evaluating the effectiveness of their re-teaching strategies	2.78	0.52	0.36	0.12	55
TEACHERS					
Thinking about your school's instructional leader(s), how would you rate their ability to do each of the following activities:					
Monitor the quality of teaching at this school	3.53	0.23 *	0.21	0.09	523
Participate in instructional planning with teachers	3.46	0.20	0.17	0.22	524
Institute concrete practices and procedures that encourage the use of student test data by teachers	3.71	0.23 *	0.22	0.10	523
Provide actionable feedback on instructional plans	3.36	0.30 **	0.27	0.04	524

Note: Total leader n = 57; total teacher n = 527. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Standard errors were clustered by school for all teacher-level analyses.

Leader frequency items response scale: 1=never, 2=once or twice a year, 3=quarterly, 4=monthly, 5=weekly or more often. Teacher ratings of leader abilities: 1=very poor, 2=poor, 3=fair, 4=good, 5=excellent.

* significant at .10; ** significant at .05.

Source: Year-one (2011-12) school leader and teacher surveys.

Table 14. Teacher Actions: Analyze Results (Teacher-reported)

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i> -value	N
TEACHERS					
<i>Conversations</i>					
How often have you had conversation with colleagues in your school about:					
Student progress in literacy and/or mathematics	4.05	0.11	0.14	0.12	522
Student achievement data	3.67	0.10	0.11	0.22	525
During CPT, how often have teachers discussed:					
Student test results	3.93	0.13	0.16	0.11	507
Students who are not meeting grade level expectations	3.82	0.09	0.11	0.25	509
How often have you reviewed interim assessment data:					
Independently	3.16	0.19 *	0.22	0.09	511
With other teachers in your grade or subject area	2.88	0.29 ‡	0.36	0.00	509
With all teachers in your schools	2.09	0.30 ‡	0.35	0.01	508
With your principal, coach, or other instructional leader	2.56	0.30 ‡	0.36	0.01	511
<i>Use</i>					
Thinking about the use of CPT, please indicate the extent to which you agree or disagree with the following:					
We use student data to inform our thinking about effective instructional practice	4.16	0.18	0.18	0.11	504
Thinking about the ways you use data during CPT, how often have you used data to:					
Identify which students need additional support	4.04	0.10	0.11	0.32	508
Identify when and what to re-teach	3.78	0.44 ‡	0.43	0.00	505
Understand if a skill should be taught or re-taught to the whole class, in small groups or with individual students	3.68	0.40 ‡	0.38	0.00	504
Group students (e.g. by skill level, or in mixed-level groups)	3.67	0.10	0.09	0.44	508
Identify new materials to supplement the school's core curriculum	3.29	0.23 **	0.20	0.04	506
Refer students for additional services	3.41	-0.05	-0.05	0.62	509
How often have you used interim assessment data to:					
Plan your instruction	3.29	0.28 ‡	0.29	0.00	506
Measure student progress toward identified goals	3.25	0.23 **	0.23	0.04	505
Adjust your teaching plans to better meet students' learning needs	3.73	-0.03	-0.04	0.72	468
Identify lessons that are not needed because students have mastered the content	3.15	0.23	0.20	0.11	494

Note: Total teacher *n* = 527. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Standard errors were clustered by school for all teacher-level analyses.

Conversations/frequency items response scale: 1=never, 2=once or twice a year, 3=monthly, 4=weekly, 5=more than once a week. Discussion/frequency items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always. Reviewed/frequency items response scale: 1=never, 2=once or twice a year, 3=once or twice a month, 4=weekly, 5=more than once a week. Use/agreement items response scale: 1=strongly disagree, 2=disagree, 3=neither agree nor disagree, 4=agree, 5=strongly agree. Use/data during CPT items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always. Use/interim assessment data response items: 1=never, 2=once or twice a year, 3=monthly, 4= weekly, 5=more than once a week.

* significant at .10; ** significant at .05; ‡ significant at .01.

Source: Year-one (2011-12) teacher survey.

Table 15. Teacher Actions: Analyze Results (Leader-reported)

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i> -value	N
LEADERS					
How many teachers would you say:					
Use student assessment data to identify students in need of instructional support	3.32	0.37	0.35	0.14	55
Use student assessment data to identify which standards students have not yet mastered	2.93	0.68 **	0.57	0.02	56

Note: Total leader n = 57. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools.

Item scale: 1=very few, 2=a little less than half, 3=about half, 4=many, 5=nearly all.

** significant at .05.

Source: Year-one (2011-12) school-leader survey.

Table 16. Teacher Actions: Backward Planning, Action Plans, and Re-teaching (Teacher-reported)

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i> -value	N
TEACHERS					
<i>Conversations</i>					
How often have you discussed instructional strategies with other teachers:					
In your school	4.09	0.13	0.16	0.20	527
In other schools in your district	2.76	-0.02	-0.02	0.87	525
Outside of your district	2.13	0.03	0.02	0.79	525
How useful do you find the opportunities you have to discuss instructional strategies with teachers:					
In your school	2.35	-0.07	-0.12	0.22	457
In other schools in your district	2.60	0.10 #	0.17	0.11	524
Outside of your district	2.19	-0.06	-0.09	0.41	337
How often have you had conversation with colleagues in your school about:					
Development of new lesson plans	3.69	0.19 *	0.20	0.07	524
During CPT, how often have teachers discussed:					
Developing lesson plans	3.13	0.35 **	0.31	0.01	508
Instructional methods or pedagogy	3.40	0.20 **	0.21	0.03	509
Thinking about the use of CPT, please indicate the extent to which you agree or disagree with the following:					
We revisit and revise our thinking about effective instructional practices	3.87	0.12	0.12	0.26	501
Our discussions influence my classroom practice	4.06	0.18 *	0.18	0.06	500
<i>Planning</i>					
When planning instruction, how often do you:					
Begin by identifying the skill or goal you hope students will master	4.45	0.06	0.08	0.38	526
Begin by identifying the state standard you hope students will master	3.99	0.12	0.12	0.15	526
Create differentiated instruction plans to meet student's individualized learning needs	3.97	0.10	0.12	0.16	526
Use a curriculum scope and sequence to design lesson or unit plans	3.96	0.13	0.14	0.18	523
Schedule re-teaching time into your lesson or unit plans	3.75	0.09	0.10	0.26	525
Schedule re-teaching time outside of regular class time	3.10	0.03	0.03	0.74	525
<i>Re-teaching</i>					
How often do you do each of the following in response to students' different learning needs:					
Teach or re-teach content to the whole class	3.73	0.22 ‡	0.29	0.00	524
Teach or re-teach content to small groups of students	3.96	0.04	0.05	0.69	522
Teach or re-teach content to individual students	3.77	-0.06	-0.08	0.48	519

Note: Total teacher n = 527. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Standard errors were clustered by school for all teacher-level analyses.

Discussion/frequency items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always. Discussions/usefulness items response scale: 1=not at all useful, 2= somewhat useful, 3=very useful. Conversations/frequency items response scale: 1=never, 2=once or twice a year, 3=monthly, 4=weekly, 5=more than once a week. Discussions/CPT items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always. Use/CPT items response scale: 1=strongly agree, 2=disagree, 3=neither agree nor disagree, 4=agree, 5=strongly agree. Planning items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always. Grouping items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always.

significant at .10 (probit model only); * significant at .10; ** significant at .05; ‡ significant at .01.

Source: Year-one (2011-12) teacher survey.

Table 17. Teacher Actions: Backward Planning, Action Plans, and Re-teaching (Leader-reported Only)

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i> -value	N
LEADERS					
How many teachers would you say:					
Provide instruction to meet individual student learning needs	3.04	-0.04	-0.03	0.89	56
Re-teach content to students who aren't successful the first time	3.25	0.29	0.25	0.29	56
Try another instructional approach when students aren't successful the first time	3.04	0.16	0.14	0.57	56
During meetings to discussion interim assessment data with teachers, how often did the following occur:					
Time was provided during the meeting for lesson planning.	2.15	1.30 ‡	1.11	0.00	55
Instructional strategies to target identified priority skills were shared.	3.11	0.70 ‡	0.60	0.01	55
School leaders communicated with teachers about the specific students or skills on which they should focus their instruction.	3.30	0.34	0.29	0.21	55

Note: Total leader n = 57. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools.

Proportion of teachers item scale: 1=very few, 2=a little less than half, 3=about half, 4=many, 5=nearly all. Data meeting frequency items scale: 1=never, 2=once in a while, 3=sometimes, 4=often, 5=almost always.

‡ significant at .01.

Source: Year-one (2011-12) school-leader survey.

Table 18. Beliefs about Interim Assessment Data & Summative Assessments

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	Marginal Effect	<i>p</i> -value	N
TEACHERS						
Please tell us the extent to which you agree with the following statements about interim assessments:						
Interim assessments keep me accountable to other teachers in my school for my students' progress	3.19	0.26 **	0.25	--	0.03	509
Interim assessments keep me accountable to my school leader(s) in my school for my students' progress	3.52	0.27 **	0.28	--	0.02	508
Interim assessments place pressure on me to raise student achievement	3.82	0.02	0.03	--	0.79	507
Administering interim assessments takes too much class time away from instruction	3.31	-0.06	-0.06	--	0.62	504
Which factors, if any, hinder you from using interim assessment data to plan or modify your instruction:						
Test data is unclear or difficult to understand	0.13	-0.03	--	-0.03	0.26	506
Test data does not clearly identify which students are struggling with specific English language arts or math skills	0.14	-0.03	--	-0.03	0.36	506
Test data does not identify which questions individual students answered incorrectly.	0.06	-0.03	--	-0.02	0.22	507
Comparison data, showing how my class performed compared with other students at my school, district or state, is not provided.	0.14	-0.10 ‡	--	-0.10	0.00	506
Test data is not aligned with state standards.	0.05	0.07 **	--	0.07	0.03	506
Test data is not aligned with the curriculum that I teach.	0.10	0.16 ‡	--	0.16	0.00	507
Test materials are inappropriate for the specific students I teach	0.22	0.02	--	0.02	0.61	506
Tests are given at the wrong time of year to be useful	0.16	-0.06	--	-0.06	0.11	506
Test data is received too late to be useful	0.29	-0.15 ‡	--	-0.15	0.00	506
School or district policies do not allow me the flexibility to adjust my instruction based on student test data.	0.25	-0.19 ‡	--	-0.19	0.00	506
End-of-year state tests only evaluate what is easy to measure, not what is important	3.28	-0.12	-0.12	--	0.21	523
LEADERS						
End-of-year state tests only evaluate what is easy to measure, not what is important	2.39	0.06	0.07	--	0.79	55

Note: Total teacher n = 527. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Standard errors were clustered by school for all teacher-level analyses.

Discussion/frequency items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always. Discussions/usefulness items response scale: 1=not at all useful,

2= somewhat useful, 3=very useful. Conversations/frequency items response scale: 1=never, 2=once or twice a year, 3=monthly, 4=weekly, 5=more than once a week.

Discussions/CPT items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always. Use/CPT items response scale: 1=strongly disagree, 2=disagree, 3=neither

agree nor disagree, 4=agree, 5=strongly agree. Planning items response scale: 1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always. Grouping items response scale:

1=never, 2=rarely, 3=sometimes, 4=often, 5=almost always.

** significant at .05; ‡ significant at .01.

Source: Year-one (2011-12) school leader and teacher surveys.

Table 19. Beliefs about Interim Assessment Data & Summative Assessments

	Control Group Mean	OLS Coeff.	Effect Size (<i>d</i>)	<i>p</i>- value	N
LEADERS					
How many teachers would you say:					
Act in a manner that communicates they are invested in improving their teaching	3.75	0.31	0.32	0.19	56
Collaborate with other teachers around instruction	3.57	0.32	0.34	0.21	56
Act in a manner that communicates that all students can learn	3.63	0.19	0.20	0.45	55
Set high expectations for students	3.68	0.09	0.10	0.71	55
TEACHERS					
How many teachers in this school:					
Feel responsible to help each other do their best	2.81	-0.05	-0.07	0.57	525
Feel responsible when students in this school fail	2.72	-0.07	-0.09	0.47	526
Are really trying to improve their teaching	2.88	-0.01	-0.01	0.93	525
Are willing to take risks to make this school better	2.69	-0.01	-0.01	0.90	523
Motivate students to learn	2.81	0.05	0.07	0.53	526

Note: Total leader n = 57; total teacher n = 527. The OLS coefficient is the result of the regression of the survey variable on treatment status with statistical controls for network and whether they belong to the pair with three schools. Standard errors were clustered by school for all teacher-level analyses. Leader items response scale: 1=very few, 2=a little less than half, 3=about half, 4=many, 5=nearly all. Teacher items response scale: 1=none, 2=some, 3=most, 4=all.

Source: Year-one (2011-12) school leader and teacher surveys.