Second-Year Results From an Efficacy Study of the Acuity Data System

Jeffrey C. Wayman, Ph.D.
Shana Shaw, Ph.D.
Vincent Cho, M.Ed.

The University of Texas at Austin

March 4, 2011

Acknowledgements

We would like to acknowledge the contributions of various individuals in the work of this report. First, we wish to thank David Seitter, Linda Bond, Dan Lewis, Mark Mantel, and Tariq Durrani of CTB/McGraw-Hill for their support and input as we collected data for this study. We would also like to thank Joe O'Reilly and Cindy Bochna of Mesa Public Schools for their generous support in providing data and answering questions.

This study is funded by CTB/McGraw-Hill. CTB/McGraw-Hill has a number of efficacy studies to evaluate the effectiveness of the Acuity system and we feel fortunate to be trusted to do this research. By agreement, CTB/McGraw-Hill has no influence or control over the results or reporting from this study, making this a fully unbiased study of Acuity. The findings and conclusions from this report are solely the responsibility of the authors.

We thank CTB/McGraw Hill for undertaking a relationship of this sort. While many vendors may be uncomfortable giving up control of the content of the evaluation, this study represents a commitment by CTB/McGraw-Hill to unbiased evaluation and the betterment of education. We commend them for their efforts and we endeavor to provide good work in return.

Executive Summary

Introduction

This study is the culmination of a two-year project examining the effects of the Acuity assessment system in Mesa Public Schools. The primary goal of this project was to identify the relationship between teacher Acuity use and student achievement. Accordingly, we endeavored to identify how teachers used Acuity, what their attitudes were toward Acuity, and how the district supported Acuity use. Toward this goal, we explored three research questions:

1. What teacher factors are associated with Acuity use?
2. Does Acuity use correlate with changes in attitudes toward Acuity?
3. Does Acuity use correlate with improvements in student achievement?

Method

We employed a mixed-methods design in conducting this study. Qualitative data were collected via interviews and focus groups. Quantitative data included a district-wide survey of data use, teacher background data, teacher Acuity use logs, student background data, and student achievement data.

The focus of the current report is our quantitative analyses. Because of the nested structure of the data (students within teachers and teachers within schools), we used Hierarchical Linear Modeling (HLM) to answer our research questions.

To answer Research Question One, we focused on two measures of Acuity use: prevalence and consistency. Prevalence of Acuity use was measured by counting the number of actions in instructional areas of the system (e.g., Reports) during the 2010 school year. Consistency of Acuity use was measured by counting the number of weeks a teacher accessed instructional areas in 2010. Independent variables included one teacher level variable (years of experience) and two school-level variables (AZ Learns rating and school level (elementary or junior high)).

To answer Research Question Two, we estimated Acuity attitudes (our dependent variable) by a nine-item survey scale that asked teachers their opinions of the Acuity system. The independent variable of interest was Acuity use, which was estimated by two measures: prevalence and consistency. Separate models were estimated for each measure of Acuity use. In each model, covariates included two teacher level variable (years of experience and prior year Acuity scale score) and two school-level variables (AZ Learns rating and school level (elementary or junior high)).

To answer Research Question Three, we used two measures of student achievement as dependent variables: 2010 state tests (AIMS) and 2010 Acuity predictive assessments. The independent variable of interest was Acuity use, which was estimated by two measures: prevalence and consistency. Separate models were estimated for each measure of Acuity use, additionally, elementary students were modeled separately from junior high students. Each of these models included student-level covariates (prior year achievement, gender, ethnicity, economically disadvantaged status) and a teacher-level covariate (years of experience).

Results

In the following sections, we highlight results pertaining to each research question.

Research Question One: What teacher factors are associated with Acuity use? At the teacher level, years of experience was not a strong predictor of Acuity use by either measure

(p= 0.169 for prevalence and p= 0.692 for consistency). School-level variables were not significantly associated with prevalence, but they were significantly associated with consistency of use: elementary teachers used Acuity 3.30 weeks more than did junior high teachers (p= 0.002).

Research Question Two: Does Acuity use correlate with changes in attitudes toward Acuity? Attitudes toward Acuity were somewhat more positive in 2010 than in 2009. However, we could not demonstrate that improved attitudes were directly associated with increased Acuity use. Neither Acuity use measure was associated with improved attitudes toward Acuity (p= 0.433 for total uses and p= 0.264 for weekly uses). Although not the focus of our research question, we note that elementary teachers demonstrated more improvement in Acuity attitudes than did junior high teachers. This effect was 0.23 points in the prevalence model (p= 0.007) and 0.22 points in the consistency of use model (p= 0.011).

Research Question Three: Does Acuity use correlate with improvements in student achievement? The effects of Acuity use on student achievement were inconsistent. In elementary schools, teacher Acuity use was demonstrated to have a statistically significant effect on AIMS reading and math scores, but these effects may not be educationally significant. For instance, prevalence of Acuity use was significantly associated with AIMS elementary reading scores (p= 0.022), but with a regression coefficient of 0.01, even a 200-action difference in Acuity use would only be associated with a two-point AIMS scale score increase. Similarly, prevalence was a significant predictor of AIMS elementary reading scores (p= 0.027), as was consistency (p= 0.021), but neither are likely to carry educational significance. Significance was not demonstrated for other elementary achievement measures, and Acuity use was not demonstrated to be significantly associated with any achievement measures in junior high schools.

We should note that larger effects were seen in our demographic student covariates of gender, ethnicity, and economic status. Thus, Acuity use is only one factor in a constellation of factors that influence student achievement growth. This constellation is notable in the stubborn persistence of these demographic factors.

Discussion

We believe it is good that MPS teachers are using Acuity more and we are optimistic that these effects may lead to stronger effects in the future. We see the results of the present study as progress, but progress that is coming too slow. We believe changes in student achievement can happen faster than have been witnessed in this and other studies of data use, and we believe these changes can be realized through systemic support of data use.

In the full report, we discuss four areas that are especially important to a systemic response in the MPS context: (1) building consistency in using Acuity, (2) embedding Acuity use in teacher work, (3) supporting the principalship, and (4) the nexus of time, collaboration, and professional learning. More specifically, we offer a few concrete suggestions that MPS may choose to address in improving Acuity and data use:

- Create activities that help build common understandings about teaching, learning, and how data may support these. We suggest identifying one specific area for discussion, such as what features educators find valuable in Acuity.
- Help principals identify a queue of questions that they can address with their teachers throughout the year. Initially, this might involve a series of principal meetings where central office personnel and others with data expertise work directly with principals

around data from their school. Soon, this could move to helping principals work with their teachers to identify issues for "the queue."

- Provide numerous district-supported simple opportunities for collaboration around data, and center them around a very specific question or topic. Aim for consistency – remember that educators should always be working on *something*.
- Consider a district change in scheduling that would explicitly free up time in the day for teachers to collaborate around Acuity.
- Look to infuse data use into the already-existing Professional Learning Communities initiative.
- Expand the idea of teacher learning from traditional "professional development" models to "professional learning" models
    - Bring professional learning to the teachers in small doses. For instance, an instructional coach might pay a 10-minute visit to a junior high math team to show them how to access data to solve a particular problem.
    - Remember that expertise does not only reside in "experts." Look to draw upon areas of expertise within many individuals. This may come from an Acuity expert, an administrator, or the teacher down the hall.
    - Avidly support collaborative reflection as the primary learning method for teachers.
- Provide extra help for junior high teachers. Our results showed that Acuity was less effective in these contexts and achievement is more variable than in elementary contexts.
- Involve CTB/McGraw-Hill in MPS efforts to broaden and deepen teacher use of Acuity. Mutual learning opportunities are there for both sides: CTB/McGraw-Hill knows their system capabilities and MPS can provide a laboratory that the vendor may not get from other sites.

Conclusion

We believe the Acuity system can be a good support for teaching and learning in Mesa Public Schools. We are optimistic that the results presented here are precursors of future results that will show Acuity use to be a major factor in student learning.

As we look to the future of MPS, we are hopeful that the district is in the early stages of building an effective information initiative. We applaud the technical infrastructure that has been built, but we know there is work yet to be done in providing the support needed to make this infrastructure educationally impactful. We hope the study we have conducted has provided a great deal of information to advance this initiative and we look forward to observing the future success of this initiative.

Introduction

This study is the culmination of a two-year project examining the effects of the Acuity assessment system in Mesa Public Schools. The primary goal of this project was to identify the relationship between teacher Acuity use and student achievement. Accordingly, we endeavored to identify how teachers used Acuity, what their attitudes were toward Acuity, and how the district supported Acuity use.

During the last ten years, research on the effective use of data has increased exponentially. Studies have considered the district as an organization and the uses of data by various roles (e.g., Copland, 2003; Datnow, Park, & Wohlstetter, 2007; Lachat & Smith, 2005; Wayman & Stringfield, 2006; Wayman, Jimerson, & Cho, in press; Young, 2006). Such research has typically focused on the conditions and practices that make data use more effective for educators, theorizing that more effective educators will result in better student achievement.

A computer system that delivers student data to teachers is one component of an effective data initiative. By delivering more data to teachers in less time, it is thought that teachers can be provided with more information on their students – and that the more they know about their students, the better teaching they will be able to provide (Wayman & Cho, 2008; Wayman, Cho, & Richards, 2010; Wayman, Stringfield, & Yakimowski, 2004). However, this premise has yet to be proved. In fact, one recent study has shown that data use is still difficult for teachers, even in the presence of adequate technology (Means, Padilla, DeBarger, & Bakia, 2009).

In the present study, we seek to provide research into the effects of a student assessment system on student achievement. We integrated two years of data, including student achievement results, weekly counts of Acuity use, and surveys of teacher attitudes. In doing so, we closely examined the Acuity use patterns of mainstream classroom teachers[1] and modeled their effects on student achievement. Consequently, the goal of this study was to examine teacher Acuity use and its effects on student achievement. Toward this goal, we explored three research questions:

1. What teacher factors are associated with Acuity use?
2. Does Acuity use correlate with changes in attitudes toward Acuity?
3. Does Acuity use correlate with improvements in student achievement?

In answering Research Questions 1 and 2, we gain a sense of the current state of Acuity use: we will know if certain types of teachers used Acuity more or less in 2010 and we will know if a second year of using Acuity has resulted in changes in attitudes toward Acuity. These answers will provide context for the main focus of our study: is there a correlation between improved student achievement and Acuity use (Research Question 3)?

The focus of this report is different than in our first-year report (Wayman, Cho, & Shaw, 2009a). Last year's report focused primarily on detailed descriptions of how teachers felt about Acuity and why they used it. Accordingly, the data used for that report was primarily qualitative and descriptive in nature. The focus for this year's report is on quantitative data – in particular, those data which identify connections between Acuity and student achievement. Thus, where last year's sample included many different kinds of teachers, this year's sample includes only mainstream teachers. We focus on students in grades four through eight because they have taken state achievement tests in each of the two years of our study.

[1] "Mainstream teachers" are teachers who are responsible for a classroom but are not responsible for special programs such as English Language Learners or Special Education.

Our report consists of four sections. First, we present the methodology for our study. Second, we offer our results and analyses. Third, we provide discussion of these results and recommendations for improvement. Finally, we offer a short conclusion section.

Method

We employed a mixed-methods design in conducting this study. Qualitative data were collected via interviews and focus groups. Quantitative data included a district-wide survey of data use, teacher background data, teacher Acuity use logs, student background data, and student achievement data.

When referring to the data educators use, we conceive of the term broadly, as any piece of information that helps teachers know more about their students. One form of "data" that is a primary focus of this study is formal achievement data: summative state achievement tests (AIMS) and interim Acuity predictive tests. But our conception of "data" also includes less-formally created data such as teacher-created benchmark assessments, tests, and quizzes. Further, we believe it is teacher judgment that brings meaning to these data.

This section outlines our methods and procedures for conducting this study. We first describe the Acuity system and the MPS context. Following this, we offer sections outlining our procedures for collecting data, the measures used, and analyses employed. Throughout this report, we refer to the 2008-2009 school year as "2009" and the 2009-2010 school year as "2010."

About Acuity

Acuity is a software program by CTB/McGraw-Hill that offers numerous functions for accessing data and standards-based instructional content. It is intended to serve educators by assessing student progress on state learning standards and determining students' readiness for state tests (CTB/McGraw-Hill, 2009). Acuity assessments may be offered online or via pencil and paper, the results of which are then accessible via the Acuity data system. These assessments target reading and mathematics in grades 3-8, as well as Algebra. When fully implemented, Acuity offers three predictive assessments to assess learning and predict progress toward state tests, four diagnostic assessments to assess learning, and a variety of reports and tools for working with these data. For instance, Acuity offers roster and individual level reporting, distracter analysis, item banks, and the ability to create and customize assessments for individual students. Our research team was trained in Acuity by CTB staff and provided access to MPS Acuity data for the purposes of becoming more familiar with the system.

There are three forms of predictive tests; the material on these tests is based on state standards. Form A is a baseline test that is given at the start of a year. In each grade, this test consists of material from the prior grade and material that is to be taught in the current grade. Form B is a mid-year test that is typically given in late Fall. It consists of material already taught, but also contains material yet to be taught. Form C is given in the early Spring and is intended to prepare students for the state test by testing all material for the current grade.

Districts are able to determine what levels of functionality are accessible to various users. For instance, central office administrators usually have access to different functions or data than principals, who have different levels of access than teachers. In this report, we focus on the instructional functions provided to MPS teachers. To evaluate these functions, we examined use logs generated by the Acuity system to track weekly educator use. We categorized instructional functions into five areas: Custom Tests, Instructional Resources, Management, Reports, and Tracking Completion Status (see *Measures* for specific information on how each of these functions were quantified).

Mesa Public Schools and Acuity Implementation

Mesa Public Schools (MPS) is located in Mesa, Arizona. The city of Mesa is a suburb of Phoenix, Arizona and has a population of approximately 450,000 (Census Bureau, 2009). For the 2010 school year, MPS enrolled about 75,000 students. Approximately 75% of MPS students identify English as their primary language at home. The district's two largest student demographic groups are white (50%) and Hispanic (38%). Roughly 62% of students are enrolled in free or reduced lunch (Mesa Public Schools, 2010). MPS has 58 elementary schools (serving grades K-6), 13 junior high schools (serving grades 7-9), and seven senior high schools (serving grades 10-12), not including other more specialized schools or academies.

MPS uses a variety of formal assessments to track student learning. The state test is one such example: *Arizona's Instrument to Measure Standards* (AIMS) is a state-mandated, criterion-referenced test, used since the 1990s. The assessment underwent extensive revision and was re-released in its current form in 2005. For grades 3-8, AIMS is administered for several days, usually in April. Students in grades 10-12 follow a different testing schedule that allows for testing in October, February and April. AIMS comprises several components, including reading, writing, and mathematics. Science is offered to fourth and eighth graders, as well as to high school biology students. Another assessment used is the *Arizona English Language Learner Assessment* (AZELLA), an English proficiency test administered to ELL students once or twice per year. The *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS) is administered to evaluate reading fluency; the DIBELS is given mostly to students in kindergarten through second grade, though it may be used for students up through the sixth grade. In addition, second and ninth graders take the complete TerraNova battery, a norm-referenced assessment, while students in grades 3-8 take an abbreviated TerraNova version that is embedded within AIMS. As a district, MPS has failed to meet Adequate Yearly Progress (AYP) accountability standards, mostly due to the reading achievement of subgroups, such as those for students classified as ELL and in special education. This failure to meet state accountability standards led district officials to search for a predictive assessment tool, the search for which culminated in the selection of Acuity.

MPS selected Acuity after an extensive process intended to ensure the contribution of many perspectives (Mesa Public Schools, 2009). Some of these strategies included participation and feedback from major user groups, an adoption committee composed of likely users of the system, and a substantial evaluation and RFP process. The district administers the Acuity predictive assessments in reading and mathematics three times per school year (in August, October, and January). These assessments are intended to help teachers track student progress toward state standards and predict performance on AIMS. The Acuity predictive tests are administered on paper, with bubble sheets used for responses. Some schools scan on-site but most tests are sent to the district for scanning.

The 2010 school year is the second year of full Acuity implementation. The system has actually been available since midway through the 2007-2008 school year, but a variety of barriers precluded complete, district-wide implementation until the 2009 school year. Even so, training during 2009 focused primarily on Acuity predictive tests. By the 2010 school year (the focus of the present study), all teachers had received training in Acuity functions relating to predictive tests, and most had received training on custom tests and instructional resources.

Procedures

Numerous types of data were collected for this study, including an online survey, district data (school, teacher, and student demographic information), weekly use logs from the Acuity system, focus groups of principals and teachers, and student achievement data. In this section, we describe our procedures for collecting data through surveys, district data, Acuity use logs, and focus groups.

Survey. The research team administered online surveys in Spring, 2010. Educators were assured in the survey invitation (and subsequent reminders) that their responses and participation would be kept confidential and that only the research team would handle these data. Each educator received an email containing a randomly-generated link to the survey; each link was electronically connected to the educator's unique MPS ID number. Participants were not allowed to leave blank responses except for demographic and open-ended questions.

Incentives were provided to boost response. Besides reminder emails, principals were encouraged to support participation rates. This was done by way of letters sent by the district office, flyers sent by the district office for posting around schools, and personal phone calls to principals. In addition, the district offered paper supplies as an incentive for reaching participation benchmarks. Each school that participated received a case of paper and all schools with at least a 60 percent response rate received more paper.

The online survey consisted of three parts: (a) A demographic section, (b) The *Survey of Educator Data Use* (Wayman, Cho, & Shaw, 2009b), and (c) The *Teacher Sense of Efficacy Scale* (Tschannen-Moran & Woolfolk-Hoy, 2001). In the demographic section, educators provided basic information about themselves, including how long they had been employed in education, what grade and subject they taught, their school, and their position (e.g., teacher, principal).

The *Survey of Educator Data Use* (SEDU) is an 81-item instrument assessing a variety of factors, including attitudes toward data use, support for data use, instructional practices, technology, and specific ways in which data are used by the respondent. This is the same survey used in our first-year report (Wayman et al., 2009a). Included in the SEDU were 14 items that asked specifically about Acuity and other MPS data systems. Measures from the *Teacher Sense of Efficacy Scale* (Tschannen-Moran & Woolfolk-Hoy, 2001) were not used for the present study.

District data. School, teacher, and student data were provided by MPS district personnel. Student data elements included gender, ethnicity, economic status, grade, and achievement scores for AIMS and Acuity in 2009 and 2010. Teacher data included years of experience in the district, grade or subject taught, and degree attainment. School demographic data included state achievement classification (Arizona Learns), and school type (elementary or junior high). Unique identifiers were available for both teachers and students that allowed linkage of students to teachers and schools.

Acuity use logs. Descriptions of teachers' use of the Acuity system, in the form of weekly use logs, were developed and provided by Acuity's vendor, CTB/McGraw-Hill. These logs reported specific actions taken and the date of the action. As such, these logs allowed us to investigate who executed which actions with the system and how often she or he performed an action. As with the teacher background data, teachers' unique identification numbers allowed linkage to student achievement data and teacher data use measures from the survey.

Focus groups. Six schools (four elementary schools and two junior high schools) were recruited as "study schools" for our previous report (Wayman et al., 2009a). As a group, these

schools were representative of the district. The same six schools agreed to participate in focus groups for the present study.

Qualitative data were collected in May, 2010 via site visits to the study schools. Two focus groups and four individual observations were conducted at each study school. Each was guided by a semi-structured protocol; whereas our previous visits to these schools focused more broadly on the use of data, our 2010 visits focused more directly on Acuity use and how it affected their instructional work. All qualitative interviews were recorded. Each participant was offered the opportunity to decline having their response recorded; none chose this option.

At each school, focus groups consisted of two types: An administrative group and a teacher group. For the administrative focus groups, principals were asked to choose three to five other participants who were knowledgeable about school data use practices. Administrative team focus groups lasted approximately one hour. Teachers were included who taught in grades 3-8 (Acuity grades), and efforts were made to speak to the same participants in 2010 that were present in 2009.

Measures and Sample Characteristics

In this section, we describe the characteristics of our sample and the measures used. Since the aim of this report is to describe the effects of Acuity on student achievement improvement over the two years of our study, we included teachers and students from grades 4-8. We included only teachers and students for whom we were able to obtain full data for both years of our study. As a result, we had 695 teachers and 19,786 students available for analysis.

In the following sections, we describe characteristics of the resulting sample and the measures used in our analyses. We begin with delimitations by teacher, then follow with sections on demographics, data use scales, use log scales, and student achievement data.

Delimitations by teacher. Since our focus was on Acuity's impact on improvements in 2010 achievement, we delimited our sample by mainstream teachers who taught in grades 4-8 during the 2010 school year. As a result, we did not include teachers primarily responsibilities included English language learners and students in special education[2]. We also did not include third grade students, because they did not take AIMS or Acuity tests in 2009. Junior high teachers were linked to a student if they were noted as the student's primary English or math teacher.

Demographics. Two school-level demographics were used: state achievement rating and type of school. Based on their AIMS performance, schools are classified into one of six categories, called "Arizona Learns" (or "AZ Learns") ratings: Failing, Underperforming, Performing, Performing Plus, Highly Performing, and Excelling. Type of school was denoted as either elementary or junior high school. We also considered using Title I status, but it was heavily confounded with AZ Learns status.

The teacher demographic variable used was years of experience in the district. This was collapsed into a four-level variable: (a) 5 years or less, (b) 6–10 years, (c) 11 – 20 years, and (d) 20 or more years. We also considered using degree attainment, but it was heavily confounded with years of experience.

Student demographics included gender, ethnicity, and economic status. For the purposes of this study, we collapsed ethnicity into three categories: Latino, Non-Latino White, and Other.

[2] Of course, the achievement of non-mainstream students is equally important. However, it carries its own set of correlates and Acuity is designed more for a mainstream environment. Consequently, we recommend a separate study to examine the achievement of these student groups.

Economic status was measured by qualification for free and reduced lunch: students were classified as Disadvantaged or Not Disadvantaged.

Most of the teachers in our sample (78%) were elementary teachers (see Table 1) and experienced teachers (see Table 2). Table 3 provides the number of teachers by their school's AZ Learns classification[3].

Tables 4-6 offer student sample descriptions for both types of school, by gender, ethnicity, and economic status.

Data use scales. Individual items on the Survey of Educator Data Use (Wayman et al., 2009b) were set on a 4-point Likert scale with response categories appropriate to the nature of the question (e.g., strongly disagree, somewhat disagree, somewhat agree, strongly agree). Five scales measuring different areas of Acuity and data use were formed from survey items. In creating each scale, responses for all the items were added and then divided by the number of items in the scale to yield an average response per scale. Scales thus ranged from one to four.

The *Acuity* scale consisted of nine items regarding the participant's attitude toward Acuity (e.g. ease of use, dependability, accuracy). The alpha reliability for this scale was 0.985.

The *Data's Effectiveness for Pedagogy* scale consisted of five items that asked about the contributions that data can make for improving educational practice (e.g., helping to plan instruction, reveal new insights, or identify learning goals). The alpha reliability for this scale was 0.919.

The *Instructional Resources* scale was an eight-item scale that asked how often participants used data for varied instructional purposes (e.g., identify students' learning needs or take actions based upon data). The alpha reliability for this scale was 0.937.

The *Principal Leadership* scale was a five-item scale that described actions that principals and assistant principals took to promote data use. This scale evaluated how school leaders led with data, encouraged teachers to use data, or created opportunities for improving data use. The alpha reliability of this scale was 0.900.

The *Support for Data Use* scale was a six-item scale assessing structures that provide support for educator data use. This included adequate preparation and professional development, as well as support from knowledgeable individuals. The alpha reliability of this scale was 0.901.

Use log data. Since the present study is focused on how Acuity helps teachers improve their practice, we focused on the instructional functions offered by Acuity. Teachers' use of instructional functions in Acuity was tracked and categorized into four areas[4] for 2010: Instructional Resources, Management, Reports, and Tracking Completion Status. Use logs generated by the Acuity system allowed us to count the number of actions within each area. We summed these to create a grand total for the year. Use logs also enabled us to count the number of weeks each teacher performed any instructional actions, to get a sense of how consistent use was throughout the year. We note that we are only measuring use in 2010: use in 2009 was so low (Wayman et al., 2009a) that no real change could be measured.

As a result, we defined three measures of Acuity use: (1) used/didn't use; (2) prevalence of use (total use actions); and (3) consistency of use (number of weeks used). We use the following sections to describe the four instructional areas in Acuity.

---

[3] MPS had no Failing schools and only one was rated as Underperforming. This school was eliminated from analysis because HLM could not estimate school effects for only one school in this category.

[4] In our previous report, use logs included a fifth area: Custom Tests. Use statistics for Custom Tests were not available for the present study. We do not believe this affects our results, because Custom Tests were not used much in 2009 (Wayman et al., 2009a) and we heard little mention of them in our 2010 focus groups.

Instructional Resources. Besides periodic assessments, Acuity offers users access to a content bank and practice assignments (instructional resources) for individual students or an entire class. Acuity offers four capabilities within the domain of instructional resources. *Assign by Skill/Standard* allows educators to assign to the entire class a resource dealing with a particular skill group or standard. *Assign by Student* allows educators to assign instructional resources to individual student(s). *Preview Instructional Resources* enables users to browse and explore the various instructional resources available. *Preview/Print* allows educators to view or print paper copies of these resources.

Management. Acuity also offers teachers limited capacity to manage administrative functions of their classes. *Manage Classes* allows teachers to manage their students' passwords, to log in as particular students, and to remove students from their roster. *Manage Student Passwords* allows teachers to manage permissions for changing passwords.

Reports. Acuity allows users to access summaries of data in the form of reports. Acuity provides seven types of reporting functions, which are often further differentiated by individual student or class level reporting. *Assessment Report* provides various summaries of Acuity periodic assessments, such as for individual students or by class. *AYP Report* provides information about performance by NCLB subgroups and their predicted performance on AIMS. *Item Analysis Report* provides information about student or class responses to particular items on Acuity periodic assessments, including information about the question, associated learning standard, correct answer, and selection of distracter responses. *Longitudinal Report* charts student progress on Acuity periodic assessments over the course of the school year. *Portfolio Report* allows teachers to look up individual students and to view summaries of their performance on various tests at once. *Roster Report* presents lists of students and their overall scores on specific periodic assessments. *Summary Report* compiles into simple figures the overall performance for a class group or student.

Tracking Completion Status. Besides the capacity to assign further instructional resources or assessments, Acuity enables teachers to track the completion status of these assignments. This can be done in two ways: by assessment or by student. The *View/Assign Status by Assessment* function allows teachers to view which assessments have been assigned and/or completed. The *View/Assign Status by Student* function allows teachers to view which assessments have been assigned to particular students and whether students have completed their assessments.

Student achievement. Student achievement was measured using state test scores (AIMS) and Acuity predictive scores. Within each, two forms of achievement were measured: mathematics and reading (called "language arts" in Acuity). Scale scores were used for each test; both are vertically equated. AIMS scores were available for both years of our study. For Acuity, we used the Predictive C forms given at the end of each school year. (Please turn to Table 12 in results for descriptive statistics on these student achievement measures.) We also note that a change in AIMS math standards for 2010 resulted in a change in the scaling. This does not affect our analyses, since we are not employing difference scores (see Analyses below).

Analyses

Because of the nested structure of our data (students nested within teachers; teachers nested within schools), Hierarchical Linear Modeling (HLM) was our primary tool of analysis. Unlike traditional methods such as ANOVA or Regression, HLM accounts for the fact that individual student responses within teachers (or teachers within schools) have some dependence

on each other. Thus, HLM gives a more accurate representation of school-to-school variance and is a more powerful method than merely modeling schools as the unit of analysis (Raudenbush & Bryk, 2002). Where helpful, further explanation was offered by our qualitative focus group data.

In the following sections, we outline the analyses we used for our study. Following the outline of our Results section, we begin with descriptive analyses, followed by a description for each of the three research questions.

Descriptive statistics. Descriptive statistics were figured looking at the sample as a whole, without accounting for the multilevel nature of our data. Three sets of descriptive statistics were constructed: (1) comparisons for the independent samples of mainstream teachers who used Acuity in 2009 and mainstream teachers who used Acuity in 2010; (2) comparisons for mainstream teachers who used Acuity in *both* 2009 and 2010; and (3) descriptions of student achievement in 2009 and 2010.

For the first and second set of comparisons (i.e., teachers), means comparisons were made between 2009 and 2010 on prevalence of Acuity use, consistency of Acuity use, and data use scales. The first set of comparisons ("independent samples") offers a snapshot look at Acuity use across the district for both years. Hypothesis tests and confidence intervals were not computed for mean differences in the independent samples because the samples are not truly independent (the samples are not defined to include the same teachers in 2009 and 2010, but many teachers are in both samples). The second set of comparisons ("paired samples") include only teachers who had valid Acuity and survey data for both 2009 and 2010. Thus, hypothesis tests and confidence intervals are presented for mean differences in these samples. For the third set of statistics, mean scale scores were computed for student achievement in both AIMS and Acuity, broken down by grade.

Research Question 1: What teacher factors are associated with Acuity use? To answer Research Question 1, we employed HLM, with teachers nested within schools. Two dependent variables were of interest: prevalence of Acuity use in 2010 and consistency of Acuity use in 2010. At the teacher level, these variables were modeled as a function of years of experience (reference category: 20+ years). The teacher-level intercept was allowed to randomly vary and was modeled as a function of two school-level variables: AZ Learns ranking (reference category: Excelling) and type of school (reference category: junior high). Significance was assessed at the 0.05 level, and 95% confidence intervals were computed for each effect.

Two models were estimated to answer Research Question 1: (1) a model that estimated associations between the prevalence of use and the independent variables; and (2) a model that estimated associations between consistency of use and the independent variables. To assess the percent of variability in each dependent variable due to schools, a null model with no independent variables was estimated; comparisons are offered between null and full models to assess the reduction in school variability due to the independent variables.

Research Question 2: Does Acuity use correlate with changes in attitudes toward Acuity? To answer Research Question 2, we employed HLM, with teachers nested within schools. Acuity attitudes were measured by the teacher's 2010 Acuity scale, which served as the dependent variable in our models. We estimated Acuity use in two ways (prevalence and consistency) and thus estimated a separate model for each type of use. In each model, years of experience was included as a the teacher level covariate (reference category: 20+ years) and the teacher-level intercept was allowed to randomly vary, modeled as a function of two school-level variables: AZ Learns ranking (reference category: Excelling) and type of school (reference

category: junior high). To isolate changes in 2010 Acuity scale, we covaried by the teacher's 2009 Acuity score. Significance was assessed at the 0.05 level, and 95% confidence intervals were computed for each effect.

Accordingly, the effects of Acuity use on Acuity attitudes were measured in with two models: The first model had had 2010 Acuity scale as the dependent variable; teacher-level variables of 2009 Acuity scale, prevalence of Acuity use in 2010, and teacher experience; and school-level variables of AZ Learns and school type. The second model was the same as the first, with consistency of use substituted for prevalence of use. Again, these full models were compared with a null model to determine the reduction of school-level variability due to teacher and school variables.

Research Question 3: Does Acuity use correlate with improvements in student achievement? To answer Research Question 3, we employed HLM, with students nested within teachers[5]. The dependent variables of interest were student scale scores on AIMS, Acuity reading assessments, and Acuity math assessments. To isolate changes in 2010 achievement, we covaried by the student's 2009 achievement score. To assess the effects of 2010 Acuity use on changes in achievement, we estimated separate models using each Acuity use measure (prevalence and consistency) as an independent variable. In each model, gender (reference category: Male), ethnicity (reference category: Other), and economic status (reference category: Disadvantaged) were included as student-level covariates, along with 2009 achievement. The student-level intercept was allowed to be randomly varying[6], modeled as a function of teacher experience (reference category: 20+) and Acuity use. Separate models were estimated for elementary and junior high schools[7]. Significance was assessed at the 0.05 level, and 95% confidence intervals were computed for each effect.

In all, 16 models help us answer Research Question 3. Models were separated into four combinations of elementary and junior high reading and math, and within each school type/subject combination, four models were estimated: (1) a model involving AIMS achievement, prevalence of use, and covariates; (2) a model involving AIMS achievement, consistency of use, and covariates; (3) a model involving Acuity achievement, prevalence of use, and covariates; (2) a model involving Acuity achievement, consistency of use, and covariates. As before, full models were compared with a null model to determine the reduction of school-level variability due to teacher and school variables.

[5] We also examined school effects as a third level. Such effects were typically not significant, so for clarity, we offer a two-level model.

[6] We explored allowing the coefficient for Acuity use to vary randomly among teachers. For most models, there was not significant random variation of this relationship. For clarity, we modeled this effect as fixed.

[7] We recognize that estimating separate models forces an interaction between type of school and all independent variables, but we believe separate models offer more clarity.

Results

In this section, we present the results of our analyses, beginning with descriptive statistics for our variables of interest. Following the descriptives section, three additional sections outline the results for each of our research questions: (1) factors associated with 2010 Acuity use, (2) the association of Acuity use with changes in Acuity attitudes, and (3) the relationship between Acuity use and student achievement.

Descriptive Statistics

In this section, we offer a descriptive look at our variables of interest. These descriptives are figured looking at the sample as a whole, without accounting for the multilevel nature of our data. Multilevel analyses are offered in sections corresponding to the three research questions.

Three sets of descriptive statistics are presented here: (1) comparisons for the independent samples of mainstream teachers who used Acuity in 2009 and mainstream teachers who used Acuity in 2010; (2) comparisons for mainstream teachers who used Acuity in *both* 2009 and 2010; and (3) descriptions of student achievement in 2009 and 2010.

Acuity use, 2009 to 2010: Independent samples. The comparisons in this section offer descriptive statistics of Acuity use through two measures: prevalence (number of total uses during the school year) and consistency (number of weeks used during the school year). Comparisons are also offered through survey scales measuring attitudes toward data. Averages are computed separately for mainstream teachers who used Acuity in 2009 and for mainstream teachers who used Acuity in 2010. While the samples are not defined to possess the same teachers in 2009 and 2010, many teachers are in both samples. Thus, the samples are not "independent" in a statistical sense and hypothesis tests of means are not appropriate. Still, the statistics offered here give an excellent snapshot description of Acuity use and data use attitudes from year to year.

Acuity use increased dramatically from 2009 to 2010, both in terms of prevalence and in terms of consistency. Whereas 70% of mainstream teachers used Acuity at least once in 2009, almost every mainstream teacher in MPS (92.5%) used Acuity in 2010 (see Table 7). Prevalence also increased greatly in the second year of Acuity implementation. In 2009, users averaged only 40 total uses per person; conversely, in 2010, users averaged 133 total uses. This increase represents an average per week increase of almost five additional Acuity uses. In addition to the increase in total Acuity uses, consistency of use also increased. The average number of weeks used climbed from 4.89 in 2009 to 10.12 in 2010 (Table 8).

The functions Acuity users accessed with greater frequency in 2010 were largely the same as those used in 2009. For instance, like those in 2009, 2010 users accessed Reports with the greatest frequency. Of the 40 uses per person in 2009, 28.63 (72%) were of Reports. Similarly, of the 133 uses per person in 2010, 100.26 (75%) were reports.

In addition to increased use, attitudes toward data use increased slightly in 2010. The most notable increase was in the Acuity scale, with an average increase of 0.13 points on a four-point scale (Table 9).

Acuity use, 2009 to 2010: Paired samples. The comparisons in this section offer descriptive statistics of Acuity use as related to prevalence and consistency, as well as the survey scales measuring attitudes toward data and data use. Averages are computed using difference scores for mainstream teachers who used Acuity in both 2009 and 2010. P-values and 95% confidence intervals are presented for each average difference.

Acuity use increased dramatically among users who accessed Acuity consecutively in 2009 and 2010. Table 10 shows that teachers averaged 105 more uses in 2010 than in 2009 (95% confidence interval: 95.23–114.89). Consistency of use also increased; Table 10 shows that teachers used Acuity an average of six more weeks during the 2010 school year (95% confidence interval: 5.85—6.77).

As measured by the survey subscales, attitudes toward data use increased slightly in 2010. Table 11 shows that the biggest increase was in the Acuity scale, with a mean difference of 0.14 (95% confidence interval: 0.08—0.21).

Student achievement in 2009 and 2010. We also explored reading and math achievement on AIMS assessments and Acuity predictive assessments for students in fourth through eighth grades[8]. Table 12 shows descriptive results for student achievement in both 2009 and 2010[9]. Note that math scores are lower in 2010. This is a result in a change in math standards, and thus, the way the tests are scaled.

Research Question One: What Teacher Factors Are Associated With Acuity Use?

In this section, we explore the results from analyses exploring the first research question: What factors were associated with mainstream teacher Acuity use in 2010? Since teachers were nested within schools, Hierarchical Linear Modeling (HLM) was used to account for this nesting and to determine which factors were associated with teachers' Acuity use (see Method for more explanation on our use of HLM).

Acuity use was measured in two dimensions: prevalence and consistency. Prevalence of Acuity use was measured by counting the number of actions in instructional areas of the system (e.g., Reports) during the 2010 school year. Consistency of Acuity use was measured by counting the number of weeks a teacher accessed instructional areas in 2010. One teacher level variable was examined for association with Acuity use: years of experience in education. Two school-level variables were examined: AZ Learns rating and school level (elementary or junior high). For each Acuity use outcome, one model estimated the percent of variability due to schools ("null model"), then a second model estimated associations between the outcome variables and the independent variables ("full model").

The null model results reveal that some variation in prevalence (16.4%) is attributable to school membership, with more variability in consistency of use being attributable to school membership (31.5%). The addition of school-level factors slightly reduced the school-level variability in both outcomes—to 16% for total uses and 25.7% for consistency of use (Table 13).

At the teacher level, years of experience was not a strong predictor of Acuity use by either measure (p= 0.169 for prevalence and p= 0.692 for consistency; see Tables 14 and 15). While experience was not a statistically significant predictor, we did observe an interesting pattern with teachers with six to ten years of experience. These teachers used Acuity more than teachers with more or less experience. In fact, the model estimates that these teachers averaged 34 more total uses per year than teachers with zero to five years of experience and confidence

[8] As discussed in Method, third grade students are not a focus of this study.

[9] The results from our multilevel modeling analyses of student achievement are presented in the section for Research Question 3.

intervals suggest this estimate may be as high as 57 uses[10]. Consistency of use showed a similar trend, but with negligible magnitude.

School-level variables were not significantly associated with prevalence (see Table 14), but they were significantly associated with consistency of use. Table 15 shows that elementary teachers used Acuity 3.30 weeks more than did junior high teachers (p= 0.002, 95% CI: 1.29—5.31). Table 15 also indicates that state rating was nearly statistically significant (p= 0.086). Schools ranking in the two lower AZ Learns categories averaged a little over two additional weeks than teachers in schools in the upper AZ Learns categories.

Research Question Two: Does Acuity Use Correlate With Changes in Attitudes Toward Acuity?

In this section, we explored the second research question: Was Acuity use associated with changes in attitudes toward Acuity? As in the previous section, HLM was used to determine associations. Acuity attitudes in 2010 was the dependent variable, measured by the Acuity scale from the survey. We were particularly interested in how 2010 Acuity use was associated with changes in Acuity attitudes, so we included 2010 Acuity use as an independent variable at the teacher level. We assessed change in Acuity attitudes by controlling for the teacher's 2009 Acuity scale score. As with research question one, Acuity use was measured through prevalence of use and consistency of use. Additional covariates included years of experience at the teacher level, state rating (AZ Learns) at the school level and school type (ES/JHS) at the school level.

Three models were estimated. As before, a null model was estimated to examine the variability in Acuity attitudes due to school membership. Next, two full models were fit, representing the two dimensions of Acuity use: prevalence and consistency.

School membership accounted for 25.2% of the variability in Acuity attitudes. The full model reduced this variability to 13.3% when defining use as total uses, and to 12.1% when defining as weeks used (see Table 16). When accounting for our control variables, Tables 17 and 18 show that neither Acuity use measure was associated with improved attitudes toward Acuity (p= 0.433 for total uses and p= 0.264 for weekly uses). Although not the focus of our research question, we note that elementary teachers demonstrated more improvement in Acuity attitudes than did junior high teachers (when controlling for experience, prior Acuity use, and state rating). This effect was 0.23 points in the prevalence model (p= 0.007) and 0.22 points in the consistency of use model (p= 0.011). Confidence intervals suggest this advantage may be as high as 0.40 in each model.

Research Question Three: Does Acuity Use Correlate With Improvements In Student Achievement?

In this section, we explore the third research question: Was Acuity use associated with student achievement improvement? As in the previous sections, HLM was used to determine associations. For this question, we fit models with students nested within teachers to analyze associations between Acuity use and student achievement outcomes. Two achievement outcomes were examined: 2010 state tests (AIMS) and 2010 Acuity predictive assessments. To assess change in achievement, we covaried by prior year achievement. As before, we defined Acuity use in two dimensions at the teacher level: prevalence and consistency. Further, we

[10] This finding should be interpreted with caution. Readers should first recall that the experience variable was not found to be statistically significant. Further, some may question whether an additional 34 uses per year is practically significant. Our sense is that this pattern is worth monitoring as MPS continues to examine its Acuity use.

included student-level covariates of gender, ethnicity, economically disadvantaged status, and included a teacher-level covariate of teaching experience. Thus, our models isolates the unique effect of 2010 Acuity use on 2010 achievement by removing effects due to school, demographic characteristics, and prior achievement.

We estimated separate models for elementary and junior high students. In the following sections, we offer four sets of achievement models: elementary reading, elementary math, junior high reading, and junior high math. Within each section, we present three models for AIMS achievement and three models for Acuity achievement: a null model, a full model for prevalence of Acuity use, and a full model for consistency of Acuity use.

Elementary reading. The null model revealed that 20.3% of variability in AIMS elementary reading scores and 21.1% of variability in Acuity elementary language arts scores is attributable to teachers. The full models for prevalence of Acuity use and consistency of use reduced teacher-level variability to 8.4% and 8.6%, respectively. Similar reductions were seen for the Acuity language arts assessments (see Table 19).

After removing the effects of covariates, prevalence of Acuity use was significantly associated with AIMS elementary reading scores (p= 0.022, see Table 20). It is unlikely this effect is educationally significant, though, because the regression coefficient is only 0.01. Thus, even substantial increases in Acuity use would still be associated with only marginal AIMS increases. For example, even a 200-action difference in Acuity use would only be associated with a two-point AIMS scale score increase. Consistency of Acuity use was not significantly associated with AIMS reading scores (p= 0.412, see Table 21).

Neither measure of Acuity use was significantly associated with 2010 elementary Acuity language arts scores. After removing the effects of covariates, neither the prevalence of Acuity use, nor the consistency of Acuity use, were significantly associated with 2010 elementary language arts achievement (p= 0.355 and p= 0.673, respectively; see Tables 22 and 23).

The focus of our investigation is on the effect of Acuity use after controlling for covariates. Still, there remain some important aspects of our covariates that illuminate the road ahead for MPS and its work to close achievement gaps. Gender, ethnicity, and economically disadvantaged status are persistently significant, after controlling for prior achievement and teacher effects. Tables 20 through 23 show that females typically outscore males by two to three points on each achievement measure, economically advantaged students score almost five points higher, and Non-Latino White students outpace minority students, sometimes by almost eight points. In isolation, these effects may not be strong, but as a constellation, they may be impactful. For example, when controlling for total number of Acuity uses by her teacher, an economically-advantaged Non-Latino White girl can be expected to experience a growth of 13.44 points more on her Acuity language arts assessment than would an economically-disadvantaged Latino boy (see Table 22).

Elementary math. The null model shows that 24.7% of variability in AIMS elementary reading scores and 26.2% of variability in Acuity elementary language arts scores was attributable to teachers. The full models for both measures of Acuity use reduced teacher-attributable variability to 21.7% for AIMS scores and 12.3% for Acuity scores (see Table 24).

Both measures of Acuity use were significantly associated with 2010 elementary AIMS math achievement, after accounting for 2009 math achievement and other covariates. Table 25 shows that the prevalence measure was a significant predictor (p= 0.027), and Table 26 demonstrates that the consistency measure was also significant (p= 0.021). Similar to elementary reading results, we question whether these effects are educationally significant. The

prevalence effect for elementary math is 0.01, as it was for reading, again implying that a 200-action difference would only be associated with a two-point difference in AIMS math scores. Similarly, the effect for consistency of use is 0.28. A substantial use difference of 10 weeks would only be associated with a three-point increase in AIMS math scores.

Neither measure of Acuity use was statistically significant in the models for Acuity math assessments. Table 27 indicates that the prevalence variable approached significance (p= 0.085), but we hold the same concerns about the practicality of this effect as we did in the AIMS analysis. Consistency of use does not approach significance (p= 0.159, see Table 28).

Recognizing that the focus of this study is not on our covariates, we will make brief mention of them, as we did in the section on elementary reading outcomes. The student demographics of ethnicity and economic status were again significant, but gender was not. The effects of these variables were not as strongly associated with math achievement as they were with reading achievement.

Junior high reading. The null model shows that 25.0% of variability in AIMS junior high reading scores was attributable to teachers, and 23.6% of variability in Acuity elementary language arts scores was attributable to teachers. For AIMS scores, the full model for prevalence reduced teacher-level variability to 3.8% and the full model for consistency of use reduced teacher-level variability in AIMS reading scores to 3.6%. For the Acuity test, the full model for prevalence reduced teacher variability to 5.7%. Teacher-level variability was reduced to 5.6% for consistency of use (see Table 29).

When controlling for our covariates, neither Acuity use measure was significantly associated with AIMS reading scores (Tables 30 & 31). However, we should note the large confidence interval associated with the consistency of use effect: with an upper bound of 1.24, it is possible that an educationally significant relationship exists in the population, but was not detected in our sample (Table 31). Similarly, neither use measure was significantly associated with Acuity language arts scores (Tables 32 and 33). As with the AIMS model, the upper bound on the consistency of use effect is quite high (1.78, see Table 33), so it is possible that an educationally significant relationship exists in the population, but was not detected in our sample.

As with elementary scores, we observed that many of our student demographic covariates were significant. Tables 30 through 33 show that ethnicity was not consistently significant, but that gender and economic status were. Females outperformed males in all models by three to four points. Students at an economic disadvantage were strikingly behind their advantaged peers on math scores: in the five-point range for AIMS scores, and in the seven point range for Acuity scores.

Junior high math. The null models show considerably more teacher-level variability in junior high math scores than in any of the prior models. Specifically, Table 34 shows that 48.6% of AIMS variability and 33.7% of Acuity variability is due to teachers. Full models for use measures reduced teacher-level AIMS variability to 13.5% for prevalence and 13.8% for consistency of use. Similarly, the full model for prevalence reduced teacher-level Acuity variability to 16.2% and the model for consistency of use reduced Acuity variability to 16.5%.

Tables 35 through 38 show that neither of our use measures was significant for either AIMS or Acuity math scores. It was difficult to precisely estimate the effects of our independent variables and covariates, probably due to the aforementioned variability in math scores. Evidence of this is seen in the unusually wide confidence intervals associated with both dimensions of Acuity use.

In the previous three sections, we noted many significant covariate effects; these effects are not consistently significant in junior high math models. The effects associated with teacher experience are notable, however, as the confidence intervals for teacher experience categories ranged up to 40 points on Acuity math.

Discussion

Our findings are clear that Acuity was used substantially more in MPS during the last school year. Almost every mainstream teacher used Acuity during 2010, averaging 5 more weeks of use and over 100 more uses than in 2009. Use was still primarily focused on areas of Acuity that provided summary reports. The district brought greater attention and support to Acuity use during the 2010 school year, and we believe this may have substantially contributed to increased use.

Increased use was fairly uniform across the district. Our models showed only minor differences: elementary schools showed an increase in consistency of use (but not in total uses), and schools that ranked lower on AZ Learns ratings used Acuity slightly more.

Attitudes toward Acuity were somewhat more positive in 2010. However, we could not demonstrate that improved attitudes were directly associated with increased Acuity use.

The effects of Acuity use on student achievement were inconsistent. In elementary schools, teacher Acuity use was demonstrated to have a statistically significant effect on AIMS reading and math scores, but these effects may not be educationally significant. Acuity use was typically not demonstrated to be significantly associated with Acuity assessments in elementary schools, or with any assessments in junior high schools. Larger effects were seen in our demographic student covariates of gender, ethnicity, and economic status. Thus, Acuity use is only one factor in a constellation of factors that influence student achievement growth. This constellation is notable in the stubborn persistence of these demographic factors.

We believe it is good that MPS teachers are using Acuity more and we are optimistic that these effects may lead to stronger effects in the future. However, we believe there are systemic conditions that must be in place to realize this benefit. In the following section, we discuss considerations for improving the effect of Acuity on student achievement. We follow this section with some concrete steps MPS may take to make Acuity more useful. We close with some concluding thoughts.

Increasing Acuity's Effects on Student Achievement

While we were unable to show clear effects of Acuity use on student achievement in MPS, the results of this report do indicate progress. We are encouraged by the large increase in use of the system and we heard some positive changes in how the system is used to improve practice. Further, while Acuity's effects on student achievement may have been modest, we believe they may be a step toward greater effects.

At this stage, the Acuity system has been implemented, but issues of practice have not been addressed. The school district provided increased training and support of Acuity during the 2010 school year, but most of this support was in terms of drawing attention to Acuity: what it is, what it can do, how data may be accessed.

Consequently, we are unsure what the continued effects of the Acuity system may be. It is possible that this year's results are one step toward greater results: we saw muted use in 2009, we saw increased use and slight achievement effects in 2010, and we may see increased use and increased achievement effects in 2011. This is in line with at least one other study that showed the effects of data use to take two or three years to manifest themselves (Stringfield et al., 2009).

On the other hand, we noted in our prior report (Wayman et al., 2009a) that teachers seemed to be assessing the "risk" in using Acuity. That is, data from both years of this study indicate that teachers are evaluating the time and effort it costs them to learn Acuity use, versus their perception of whether Acuity can give them enough new information about their students.

Without better support, we are unsure whether teachers will judge Acuity as worth their risk for next year and beyond – and thus, whether they will keep using it.

We see the results of the present study as progress, but progress that is coming too slow. We believe changes in student achievement can happen faster than have been witnessed in this and other studies of data use, and we believe these changes can be realized through systemic support of data use. We expand on these ideas in the following sections. We set the stage by offering a short perspective on appropriate systemic support. Following this, we offer four areas of discussion pertinent to a systemic response: (1) building consistency in using Acuity, (2) embedding Acuity use in teacher work, (3) supporting the principalship, and (4) the nexus of time, collaboration, and professional learning.

Systemic support for Acuity. Researchers have suggested that effective data use is best accomplished by systemic, district-wide support of the data initiative (Datnow et al, 2007; Supovitz, 2006; Wayman et al., 2010). In the present study, we observed support provided by the MPS central office, but this support was largely due to the efforts of one central office administrator. It is our belief that if MPS wants to maximize its Acuity investment, it must launch an intentional, systemic plan in support of Acuity use. Specifically, we believe the district must address problems surrounding the consistent use of Acuity within a teacher's natural workday, supporting principals in leading faculty to use data, and providing adequate time, supporting professional learning, and opportunities for collaboration. In the following sections, we outline important considerations regarding these issues.

Building consistency in using Acuity. On average, users accessed Acuity about 11 weeks out of the school year. While this represented a large increase over the previous year, it still represented less than a third of the total weeks in the school year. This is consistent with the descriptions educators gave of their use: Acuity was typically used for month-to-month issues such as dealing with instruction in large chunks of time (e.g., quarters, semesters, years), addressing groups of struggling students, and improving proficiency ratings.

Accordingly, it is no surprise that most instructional uses of Acuity involved accessing Reports: if Acuity is to be used for month-to-month issues and if teachers have to carve out time to use it, then the fast overview provided by Acuity reports are a perfect fit. But if these continue to be the primary uses of Acuity, it is very possible that Acuity use has already reached a ceiling.

We envision a different type of use. As our data shows, student achievement is a constellation of effects – and thus demands a constellation of data (e.g., Ingram, Louis, & Schroeder, 2004; Supovitz & Klein, 2003; Wayman & Stringfield, 2006). Planning for month-to-month issues is an important component of this constellation, but examining day-to-day learning issues is also critical. MPS seems to have solved "month-to-month" use; we believe MPS must now endeavor to promote use in a day-to-day fashion.

To realize this type of use, MPS educators can take better advantage of the full capabilities offered by the Acuity system. For instance, Acuity offers the ability to assign extra practice to students based on a particular skill or standard. It also offers a content bank for these purposes. And, the teacher can write her or his own custom tests to assign to students or classes[11].

In our previous report (Wayman et al., 2009a), we noted that many teachers were eager to use features such as this, but many barriers precluded them. The primary barrier was time –

[11] We note that teachers also reported these areas of Acuity cumbersome and time-consuming to use. Our own exploration of the system suggests this to be partly true, but that these issues can be overcome. MPS should stay true to the spirit of this recommendation, but should watch out for pitfalls that may arise.

many teachers were unable to fit these tasks into their natural workday. In the following section, we discuss the importance of embedding daily Acuity use into teacher work.

Embedding Acuity use in teacher work. In our previous report (Wayman et al., 2009a), we noted that teachers often did not use Acuity because it did not fit into their workday. As a consequence, effective data use practices – such as collaboration, reflection, and working toward a clear goal – were hard for teachers to realize. An initiative that was successful in spurring this work was one we called the "principal homework assignment." In this initiative, principals were sent a problem to work on; this problem was related to their school's context and dealt with Acuity data. Many principals involved their faculties in solving this problem, forming groups that looked at the assigned Acuity data. Many educators spoke positively of this work, and their comments were highlighted by collaboration, learning, and better knowledge of student needs.

Unfortunately, this event was an isolated occurrence; we did not hear many other examples of how teachers were using Acuity in everyday work. Again, time was cited as an important reason not to use Acuity, and the district had made no moves to provide such time. Collaboration opportunities remained few. And training mostly involved ways to access data from the system. We believe this consistent lack of everyday support is probably creating an Acuity culture that supports punctuated use of Acuity with a month-to-month view, rather than the day-to-day view that is most important to student learning.

We recognize that day-to-day use of data is a difficult goal to achieve. However, we also know that research suggests it will be far less difficult if effective district supports are intentionally implemented (Datnow et al, 2007; Supovitz, 2006; Wayman et al., 2010). In the following sections, we discuss two such supports we have found to be immediately important in MPS: issues relating to the principalship, and the interplay of time, collaboration, and professional learning.

Supporting principal leadership for data use. When we visited our study schools, we found many of our principals making concerted efforts to further Acuity use. Many explicitly stated the importance of Acuity in helping them and their staff know more about student learning needs. Some of our principals were also adept in identifying the primary learning issues that Acuity could inform.

Unfortunately, nearly all of the initiatives set forth by our principals dealt with month-to-month issues; we found our principals to be unsupported in facilitating the sorts of day-to-day use of Acuity that we believe is important. The district can provide this support in many ways, none of which would represent a major investment. For example, we believe most principals would benefit from help in identifying a queue of specific, short-term problems to address ongoing (we found many of our principals adept at identifying big picture issues, but less so at identifying day-to-day problems). Also, our data is clear that MPS educators do not share a common language around student learning and the use of data; principals could be supported in facilitating these conversations within their schools.

Perhaps the greatest problem faced by principals is how to provide their teachers with time to use data, time and structure to collaborate, and sufficient professional learning opportunities to advance their skills. We see these three issues as intertwined and we discuss this nexus in the following section.

Time, professional learning, and collaboration. We noted that MPS teachers have no structured time to use Acuity or to collaborate, and the training they receive is focused on how to

use the Acuity system. In contrast, we suggest that time, collaboration and professional learning should all be handled together (Wayman et al., in press).

Considering time, professional learning, and collaboration as one requires a different conception of professional learning than is currently shown in MPS. MPS seems to have solved the problem of getting educators to use Acuity. What remains is what to do after the data are accessed. Accordingly, Wayman et al. (in press) suggest that professional learning should be focused on changes in practice. The authors also suggest that such learning is best built collaboratively, in small pieces. Finally, Wayman et al. (in press) suggest that educators may learn from a variety of different "experts" – not just central office staff or instructional coaches, but also each other, in collaborative sessions.

Consequently, combining these three issues can be efficient for MPS. Such a structure would demand few resources other than offering time – once time is offered for teachers to work collaboratively, they can learn from each other and can target areas where they need further support.

A successful implementation of this was seen in "principal homework assignment" (see "Embedding Acuity use in teacher work" above). In using their time to work on this assignment, teachers reported enjoying their collaboration in part because they learned from each other while discussing their craft. This was only one occurrence, though. MPS must search to replicate this environment in daily work throughout the school year. In the following section, we will discuss some concrete suggestions of actions MPS may take to support this and other initiatives.

Moving Forward: Suggestions for MPS

In this section, we offer some concrete suggestions for how MPS may improve district-wide data use with Acuity. Our suggestions are based on the knowledge we have gained in the two years of this study, our own research in other contexts, and research of other scholars.

There are some tenets that should be followed in each solution. First, in every solution, look to stay small and targeted. That is, look for initiatives that have clear goals that can be achieved in a short amount of time. Second, stay consistent. Avoid any down time – once one question is explored, have another ready to go. Third, district educators must see that the superintendent and other central office administrators all support this initiative. We suggest an intentional statement from the superintendent's office – both in words and action[12] – that data use is an ethic in MPS. Fourth, it may not matter exactly *what* is addressed, but that *something* is being addressed. It is probably more important to be in the ballpark than to hit a home run.

In the following bullets, we offer a few concrete suggestions from which MPS may wish to choose, amend, or build upon. In implementing these, we suggest that MPS may wish to choose a small group of pilot schools to work with during the Spring 2011 semester, then look to roll out these supports to all schools for the 2011-2012 school year.

To improve data and Acuity use, MPS may choose to:

- Create activities that help build common understandings about teaching, learning, and how data may support these. We suggest identifying one specific area for discussion, such as what features educators find valuable in Acuity. (More can be found in Wayman et al. (2010) and Wayman et al. (in press)).

[12] Early, unpublished results from a study we are conducting in other districts have shown the importance of the presence of the superintendent. For example, by merely attending meetings on the data initiative, one superintendent has sent the message that she believes this initiative is important. We recognize that these results are currently anecdotal, but we believe they are important.

- Help principals identify a queue of questions that they can address with their teachers throughout the year. Initially, this might involve a series of principal meetings where central office personnel and others with data expertise work directly with principals around data from their school. Soon, this could move to helping principals work with their teachers to identify issues for "the queue."
- Provide numerous district-supported simple opportunities for collaboration around data, and center them around a very specific question or topic. The "principal homework assignment" is a good example (see "Embedding Acuity use in teacher work" above). Aim for consistency – remember that educators should always be working on *something*.
- Consider a district change in scheduling that would explicitly free up time in the day for teachers to collaborate around Acuity.
- Look to infuse data use into the already-existing Professional Learning Communities initiative.
- Expand the idea of teacher learning from traditional "professional development" models to "professional learning" models (Wayman et al., in press).
    - Bring professional learning to the teachers in small doses. For instance, an instructional coach might pay a 10-minute visit to a junior high math team to show them how to access data to solve a particular problem.
    - Remember that expertise does not only reside in "experts." Look to draw upon areas of expertise within many individuals. This may come from an Acuity expert, an administrator, or the teacher down the hall.
    - Avidly support collaborative reflection as the primary learning method for teachers.
- Provide extra help for junior high teachers. Our results showed that Acuity was less effective in these contexts and achievement is more variable than in elementary contexts.
- Involve CTB/McGraw-Hill in MPS efforts to broaden and deepen teacher use of Acuity. Mutual learning opportunities are there for both sides: CTB/McGraw-Hill knows their system capabilities and MPS can provide a laboratory that the vendor may not get from other sites.

Conclusion

Overall, we believe the Acuity system can be a good support for teaching and learning in Mesa Public Schools. Acuity provides interim information about student learning that can support day-to-day pedagogy. In the present study, we demonstrated that Acuity use can be a positive part of a collection of factors that influence student learning. We are optimistic that these results are precursors of future results that will show Acuity use to be a major factor in student learning.

As we look to the future of MPS, we are hopeful that the district is in the early stages of building an effective information initiative. A good foundation is already in place: As this report is submitted, Acuity use is in its third year, and the TetraData data warehouse is in its first year[13]. We applaud the technical infrastructure that has been built, but we know there is work yet to be done. To make effective use of the data provided by these systems, the district must provide better supports for teachers and principals in terms of time and training. And it must provide opportunities for these educators to work together and talk through solutions to these problems.

We are optimistic about the early stages of the MPS data initiative. We hope the study we have conducted has provided a great deal of information to advance this initiative. We look forward to observing – and hopefully, participating in – the future success of this initiative.

[13] TetraData integrates data from many different systems such as Acuity, gradebooks, and others, for one-stop data access.

References

Census Bureau. (2009). Mesa, Arizona. US Census Bureau. Retrieved November 28, 2010, from http://quickfacts.census.gov/qfd/states/04/0446000.html.

Copland, M. A. (2003). Leadership of inquiry: Building and sustaining capacity for school improvement. *Educational Evaluation and Policy Analysis, 25*(4), 375-395.

CTB/McGraw-Hill. (2009). Acuity Formative Assessments from CTB/McGraw-Hill. Retrieved November 28, 2010, from http://www.acuityforschool.com.

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data: How high-performing school systems use data to improve instruction for elementary students.* Los Angeles: University of Southern California, Rossier School of Education, Center on Educational Governance.

Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record, 106*, 1258–1287.

Lachat, M. A., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed At Risk, 10*(3), 333–349.

Means, B., Padilla, C., DeBarger, A., and Bakia, M. (2009). Implementing data-informed decision-making in schools: Teacher access, supports and use. Washington D.C.: U.S. Department of Education Office of Planning, Evaluation and Policy Development. Retrieved March 3, 2011, from http://www.ed.gov/rschstat/eval/tech/data-informed-decision/data-informed-decision.doc.

Mesa Public Schools. (2009, January). Acuity in Mesa Public Schools. MPS educational technology and training. Retrieved January 21, 2009, from http://www2.mpsaz.org/edtech/acuity_turnleaf.

Mesa Public Schools. (2010, February). *Mesa Public Schools Demographics*. Retrieved November 28, 2010, from http://www.mpsaz.org/research/parents/demographics/files/demographics_feb_2010.pdf

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage

Supovitz, J.A. (2006). *The case for district-based reform: Leading, building, and sustaining school improvement*. Cambridge, MA: Harvard Education Press.

Supovitz, J. A., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement*. Philadelphia: Consortium for Policy Research in Education.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783–805.

Wayman, J. C., & Cho, V. (2008). Preparing educators to effectively use student data systems. In Kowalski, T. J., & Lasley, T. J. (Eds.), *Handbook on data-based decision-making in education* (pp. 89-104). New York: Routledge.

Wayman, J. C., Cho, V., & Richards, M. (2010). Student data systems and their use for educational improvement. In McGaw, B., Peterson, P., & Baker, E. (Eds.) *The International Encyclopedia of Education* (Vol. 8, pp. 14-20). London: Elsevier.

Wayman, J. C., Cho, V., & Shaw, S. M. (2009a). *First-year results from an efficacy study of the Acuity data system.* Austin: The University of Texas. Retrieved March 3, 2011, from http://edadmin.edb.utexas.edu/datause/publications.htm.

Wayman, J. C., Cho, V., & Shaw, S. (2009b). *Survey of Educator Data Use.* Unpublished document.

Wayman, J. C., Cho, V., Jimerson, J. B., & Snodgrass Rangel, V. W. (2010, May). *The Data-Informed District: A systemic approach to educational data use.* Paper presented at the 2010 Annual Meeting of the American Educational Research Association, Denver CO.

Wayman, J. C., Jimerson, J. B., & Cho, V. (in press). Organizational considerations in establishing the Data-Informed District. *School Effectiveness and School Improvement.*

Wayman, J. C. & Stringfield, S. (2006). Technology-supported involvement of entire faculties in examination of student data for instructional improvement. *American Journal of Education 112* (4), 549-571.

Wayman, J. C., Stringfield, S., & Yakimowski, M. (2004). *Software enabling school improvement through analysis of student data* (CRESPAR Technical Report No. 67). Baltimore: Johns Hopkins University.

Young, V. M. (2006). Teachers' use of data: Loose coupling, agenda setting, and team norms. *American Journal of Education, 112*(4), 521–548.

Tables

Table 1

Number and Percent of Teachers, by Level Taught

Type of School	*n*	*%*
Elementary School	541	77.8
Junior High School	154	22.2
Total	695	

Table 2

Number and Percent of Teachers, by Experience

Experience Category	*n*	*%*
0 - 5 years	112	16.1
6 - 10 years	135	19.4
11 - 19 years	233	33.5
20 or more years	215	30.9
Total	695	

Table 3

Number and Percent of Teachers, by State Achievement Classification (AZ Learns)

AZ Learns Classification	*n*	*%*
Performing (17 schools)	136	19.6
Performing Plus (26 schools)	236	34.0
Highly Performing (16 schools)	172	24.7
Excelling (13 schools)	151	21.7
Total	695	

Table 4

Number and Percent of Students, by Gender and School Level

Gender	ES *n (%)*	JHS *n (%)*
Female	5310 (51%)	4901 (52%)
Male	5070 (49%)	4505 (48%)
Total	10380	9406

Table 5

Number and Percent of Students, by Ethnicity and School Level

Ethnicity	ES *n (%)*	JHS *n (%)*
Non-Latino White	5657 (55%)	5466 (58%)
Latino	3380 (33%)	2860 (30%)
Other	1343 (13%)	1080 (12%)
Total	10380	9406

Table 6

Number and Percent of Students, by Economic Status and School Level

Economic Status	ES *n (%)*	JHS *n (%)*
Not Disadvantaged	5413 (50%)	5109 (53%)
Disadvantaged	5449 (50%)	4505 (47%)
Total	10380	9406

Table 7

Acuity Use in 2009 and 2010, All Mainstream Teachers

	2009 *n (%)*	2010 *n (%)*
No	230 (29.6%)	52 (7.5%)
Yes	546 (70.4%)	643 (92.5%)
Total	776	695

Table 8

Average Acuity Use in 2009 and 2010, All Mainstream Teachers

Actions	2009	2010	Difference
Total use actions	39.73	128.53	88.81
Number of weeks used	4.89	10.12	5.23
Uses per week	8.12	12.70	4.58

Note. $n(2009) = 546$; $n(2010) = 643$

Table 9

Survey Scale Averages in 2009 and 2010, All Mainstream Teachers

	2009	2009	2010	2010	
Factor	*n*	Average	*n*	Average	Mean difference
Acuity	375	2.82	393	2.95	0.13
Data Effectiveness for Pedagogy	431	3.35	450	3.41	0.05
Instructional Uses of Data	431	2.37	450	2.35	-0.02
Principal Leadership	431	3.13	450	3.16	0.04
Support	431	2.79	450	2.86	0.08

Note. Acuity scale figured only for survey respondents who used Acuity during that year.

Table 10

Average Acuity use in 2009 and 2010, Paired Differences

	2009	2010				
	Average	Average	Mean Difference	95% Confidence Interval		*p*
Total use actions	37.70	142.76	105.06	95.23	114.89	0.000
Number of weeks used	4.83	11.14	6.31	5.85	6.77	0.000
Uses per week	7.80	12.81				

Note. *n*= 488

Table 11

Survey Scale Averages in 2009 and 2010, Paired Differences

	2009	2010				
	Average	Average	Mean Difference	95% Confidence Interval		*p*
Acuity	2.80	2.94	0.14	0.08	0.21	0.000
Data Effectiveness for Pedagogy	3.32	3.43	0.11	0.05	0.17	0.000
Instructional Uses of Data	2.33	2.37	0.05	-0.03	0.13	0.235
Principal Leadership	3.11	3.17	0.06	-0.01	0.13	0.078
Support	2.79	2.86	0.07	0.00	0.14	0.058

Note. Acuity scale figured only for survey respondents who used Acuity both years.

Note. Acuity *n*= 245; other scales *n*= 299

Table 12a

Student Average Achievement in Reading/Language Arts for AIMS and Acuity tests

Grade		2009 Reading (AIMS)	2010 Reading (AIMS)	2009 Lang Arts (Acuity)	2010 Lang Arts (Acuity)
4	Mean (*SD*)	477.07 (45.04)	494.96 (46.78)	448.77 (58.92)	483.86 (46.92)
	N	2990	3183	2618	2857
5	Mean (*SD*)	491.16 (43.06)	507.91 (38.62)	478.22 (47.76)	503.12 (46.39)
	N	3175	3347	2894	3100
6	Mean (*SD*)	507.2 (41.12)	522.1 (36.32)	498.38 (49.14)	524.82 (50.82)
	N	3245	3424	2983	3165
7	Mean (*SD*)	521.15 (42.69)	540.59 (43.69)	528.83 (55.36)	551.34 (62.52)
	N	4056	4410	3845	4245
8	Mean (*SD*)	534.77 (44.60)	547.3 (48.94)	552.03 (64.67)	571.9 (61.65)
	N	4689	4896	4309	4731

Table 12b

Student Average Achievement in Math for AIMS and Acuity tests

Grade		Math (AIMS)		Math (Acuity)	
		2009	2010	2009	2010
4	Mean (*SD*)	467.03 (46.34)	393.1 (41.53)	480.33 (46.38)	443.53 (45.10)
	N	2989	3182	2860	2628
5	Mean (*SD*)	505.83 (50.58)	402.94 (44.04)	503.82 (50.61)	476.66 (41.95)
	N	3177	3346	3099	2886
6	Mean (*SD*)	524.29 (51.38)	423.92 (45.20)	536.08 (62.17)	496.94 (47.21)
	N	3245	3425	3158	2968
7	Mean (*SD*)	553.47 (57.19)	444.22 (51.17)	563.58 (60.24)	539.28 (59.96)
	N	4057	4408	4294	3848
8	Mean (*SD*)	581.73 (54.87)	455.14 (44.16)	573.78 (65.72)	567.98 (55.58)
	N	4688	4896	4070	4368

Table 13

School-Level Variability in Acuity Use

Actions	Null Model	Full Model
Prevalence of use	16.4%	16.0%
Consistency of use	31.5%	25.7%

Note. *n*= 643

Table 14

Regression Model for Prevalence of Acuity Use.

Teacher-Level Variables					
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Years of Experience	0 - 5 years	-18.81	-46.24	8.62	0.169
	6 - 10 years	14.95	-11.25	41.16	
	11 - 19 years	-2.78	-24.69	19.14	
School-Level Variables					
Factor	Category	Mean Difference	95% Confidence Interval		*p*
AZ Learns	Performing	15.16	-33.31	63.63	0.272
	Performing Plus	18.63	-25.48	62.74	
	Highly Performing	-17.34	-64.89	30.21	
Elementary School		12.99	-23.36	49.34	0.478

Note. *n*= 643
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Years of Experience is 20+, AZ Learns is "Excelling;" "Underperforming" is not included in the model due to lack of underperforming schools, and Elementary School is Junior High School.

Table 15

Regression Model for Consistency of Acuity Use

		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Years of Experience	0 - 5 years	0.00	-1.25	1.26	0.682
	6 - 10 years	0.69	-0.52	1.89	
	11 - 19 years	0.13	-0.87	1.13	
		School-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
AZ Learns	Performing	2.42	-0.27	5.12	0.086
	Performing Plus	2.16	-0.30	4.62	
	Highly Performing	0.11	-2.55	2.77	
Elementary School		3.30	1.29	5.31	0.002

Note. *n*= 643
Note. Consistency of Acuity use is the number of weeks used Acuity in 2010.
Note. Reference category for Years of Experience is 20+, AZ Learns is "Excelling;" "Underperforming" is not included in the model due to lack of underperforming schools, and Elementary School is Junior High School.

Table 16

School-Level Variability in Acuity Attitudes

	Null Model	Full Model (Total Use)	Full Model (Consistency)
School variability	25.2%	13.3%	12.1%

Note. *n*= 245

Table 17

Regression Model Associating Prevalence of Acuity Use with Attitudes Toward Acuity

		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.00	0.00	0.00	0.433
2009 Acuity Attitude Scale		0.29	0.19	0.38	0.000
Years of Experience	0 - 5 years	0.00	-0.17	0.17	0.92
	6 - 10 years	-0.05	-0.20	0.10	
	11 - 19 years	-0.02	-0.14	0.10	
		School-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
AZ Learns	Performing	0.20	-0.03	0.43	0.183
	Performing Plus	0.19	-0.01	0.39	
	Highly Performing	0.08	-0.12	0.29	
Elementary School		0.23	0.07	0.40	0.007

Note. *n*= 245
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Years of Experience is 20+, AZ Learns is "Excelling;" "Underperforming" is not included in the model due to lack of underperforming schools, and Elementary School is Junior High School.

Table 18

Regression Model Associating Consistency of Acuity Use with Attitudes Toward Acuity

Teacher-Level Variables					
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Consistency of Acuity		0.01	0.00	0.01	0.264
2009 Acuity Attitude Scale		0.29	0.19	0.38	0.000
Years of Experience	0 - 5 years	0.01	-0.16	0.17	0.865
	6 - 10 years	-0.06	-0.21	0.09	
	11 - 19 years	-0.01	-0.13	0.11	
School-Level Variables					
Factor	Category	Mean Difference	95% Confidence Interval		*p*
AZ Learns	Performing	0.17	-0.06	0.40	0.311
	Performing Plus	0.17	-0.03	0.37	
	Highly Performing	0.10	-0.11	0.30	
Elementary School		0.22	0.05	0.39	0.011

Note. *n*= 245
Note. Consistency of Acuity use is the number of weeks used Acuity in 2010.
Note. Reference category for Years of Experience is 20+, AZ Learns is "Excelling;" "Underperforming" is not included in the model due to lack of underperforming schools, and Elementary School is Junior High School.

Table 19

Teacher-Level Variability in Elementary Student Reading Scores

Test	Null Model	Full Model (Total Use)	Full Model (Consistency)
AIMS	20.3%	8.4%	8.6%
Acuity predictive tests	21.1%	8.3%	8.4%

Note. *n*(AIMS)= 8791; *n*(Acuity)= 7966

Table 20

Regression Model Associating Elementary AIMS Reading Scores with Prevalence of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		2.23	1.09	3.36	0.000
Ethnicity	Non-Latino White	4.42	2.45	6.39	0.000
	Latino	1.12	-0.87	3.11	
Economically Disadvantaged		4.84	3.47	6.22	0.000
Prior Year Reading Score		0.67	0.65	0.68	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.01	0.00	0.02	0.022
Years of Experience	0 - 5 years	-0.39	-3.41	2.63	0.666
	6 - 10 years	-1.26	-4.06	1.55	
	11 - 19 years	-1.40	-3.82	1.03	

Note. *n*= 8791
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 21

Regression Model Associating Elementary AIMS Reading Scores with Consistency of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		2.23	1.10	3.37	0.000
Ethnicity	Non-Latino White	4.41	2.44	6.38	0.000
	Latino	1.10	-0.90	3.09	
Economically Disadvantaged		4.82	3.45	6.20	0.000
Prior Year Reading Score		0.67	0.65	0.68	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.07	-0.10	0.24	0.412
Years of Experience	0 - 5 years	-0.48	-3.52	2.57	0.692
	6 - 10 years	-0.95	-3.78	1.87	
	11 - 19 years	-1.46	-3.90	0.98	

Note. *n*=8791
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 22

Regression Model Associating Elementary Acuity Language Arts Scores with Prevalence of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		2.99	1.43	4.55	0.000
Ethnicity	Non-Latino White	7.78	5.05	10.50	0.000
	Latino	1.98	-0.78	4.73	
Economically Disadvantaged		4.65	2.77	6.53	0.000
Prior Year Lang. Arts Score		0.56	0.55	0.58	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.01	-0.01	0.02	0.355
Years of Experience	0 - 5 years	-2.02	-6.11	2.07	0.675
	6 - 10 years	-1.88	-5.70	1.94	
	11 - 19 years	-1.52	-4.82	1.78	

Note. n= 7966
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 23

Regression Model Associating Elementary Acuity Language Arts Scores with Consistency of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		2.99	1.43	4.56	0.000
Ethnicity	Non-Latino White	7.77	5.05	10.50	0.000
	Latino	1.96	-0.79	4.72	
Economically Disadvantaged		4.64	2.76	6.53	0.000
Prior Year Lang. Arts Score		0.56	0.55	0.58	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.05	-0.18	0.27	0.673
Years of Experience	0 - 5 years	-2.08	-6.18	2.03	0.683
	6 - 10 years	-1.73	-5.56	2.09	
	11 - 19 years	-1.56	-4.87	1.74	

Note. n= 7966
Note. Consistency of Acuity use is the number of weeks used last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 24

Teacher-Level Variability in Elementary Student Math Scores

Test	Null Model	Full Model (Total Use)	Full Model (Consistency)
AIMS	24.7%	21.7%	21.7%
Acuity predictive tests	26.2%	12.3%	12.3%

Note. *n*(AIMS)= 8232; *n*(Acuity)= 7942

Table 25

Regression Model Associating Elementary AIMS Math Scores with Prevalence of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		0.21	-0.88	1.29	0.709
Ethnicity	Non-Latino White	3.67	1.78	5.57	0.000
	Latino	0.59	-1.33	2.50	
Economically Disadvantaged		2.79	1.46	4.12	0.000
Prior Year Math Score		0.65	0.64	0.67	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.01	0.00	0.03	0.027
Years of Experience	0 - 5 years	2.37	-2.00	6.74	0.617
	6 - 10 years	-0.55	-4.63	3.53	
	11 - 19 years	1.10	-2.41	4.62	

Note. n= 8232
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 26

Regression Model Associating Elementary AIMS Math Scores with Consistency of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		0.21	-0.87	1.29	0.703
Ethnicity	Non-Latino White	3.67	1.78	5.56	0.000
	Latino	0.55	-1.36	2.47	
Economically Disadvantaged		2.81	1.48	4.14	0.000
Prior Year Math Score		0.65	0.64	0.67	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.28	0.04	0.52	0.021
Years of Experience	0 - 5 years	2.05	-2.33	6.42	0.708
	6 - 10 years	-0.57	-4.64	3.51	
	11 - 19 years	0.88	-2.63	4.40	

Note. n= 8232
Note. Consistency of Acuity use is the number of weeks used last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 27

Regression Model Associating Elementary Acuity Math Scores with Prevalence of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		-1.74	-3.51	0.04	0.055
Ethnicity	Non-Latino White	2.32	-0.77	5.42	0.002
	Latino	-1.79	-4.92	1.34	
Economically Disadvantaged		3.93	1.77	6.09	0.000
Prior Year Math Score		0.72	0.70	0.74	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.01	0.00	0.03	0.085
Years of Experience	0 - 5 years	4.18	-1.25	9.62	0.348
	6 - 10 years	-0.36	-5.41	4.70	
	11 - 19 years	2.25	-2.10	6.61	

Note. n= 7942
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 28

Regression Model Associating Elementary Acuity Math Scores with Consistency of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		-1.73	-3.50	0.04	0.055
Ethnicity	Non-Latino White	2.32	-0.78	5.41	0.002
	Latino	-1.84	-4.97	1.29	
Economically Disadvantaged		3.95	1.79	6.11	0.000
Prior Year Math Score		0.72	0.70	0.74	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.21	-0.08	0.51	0.159
Years of Experience	0 - 5 years	3.96	-1.49	9.41	0.413
	6 - 10 years	-0.24	-5.30	4.82	
	11 - 19 years	2.09	-2.28	6.45	

Note. *n*= 7942
Note. Consistency of Acuity use is the number of weeks used last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 29

Teacher-Level Variability in Junior High Student Reading Scores

Test	Null Model	Full Model (Total Use)	Full Model (Consistency)
AIMS	25.0%	3.8%	3.6%
Acuity predictive tests	23.6%	5.7%	5.6%

Note. *n*(AIMS)= 4107; *n*(Acuity)= 4531

Table 30

Regression Model Associating Junior High AIMS Reading Scores with Prevalence of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		3.97	2.08	5.85	0.000
Ethnicity	Non-Latino White	2.52	-0.79	5.83	0.133
	Latino	0.36	-3.10	3.82	
Economically Disadvantaged		5.05	2.83	7.27	0.000
Prior Year Reading Score		0.71	0.68	0.73	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.01	-0.02	0.05	0.488
Years of Experience	0 - 5 years	-4.96	-12.61	2.69	0.001
	6 - 10 years	1.83	-6.86	10.52	
	11 - 19 years	5.21	-2.19	12.61	

Note. *n*= 4107
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 31

Regression Model Associating Junior High AIMS Reading Scores with Consistency of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		3.97	2.08	5.85	0.000
Ethnicity	Non-Latino White	2.50	-0.80	5.81	0.135
	Latino	0.35	-3.11	3.81	
Economically Disadvantaged		5.03	2.81	7.25	0.000
Prior Year Reading Score		0.71	0.68	0.73	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.52	-0.19	1.24	0.145
Years of Experience	0 - 5 years	-4.36	-11.44	2.72	0.001
	6 - 10 years	2.58	-5.61	10.77	
	11 - 19 years	5.84	-1.03	12.71	

Note. *n*=4107
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 32

Regression Model Associating Junior High Acuity Language Arts Scores with Prevalence of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		3.22	0.34	6.10	0.028
Ethnicity	Non-Latino White	7.25	2.24	12.27	0.013
	Latino	4.06	-1.18	9.30	
Economically Disadvantaged		7.50	4.07	10.93	0.000
Prior Year Lang. Arts Score		0.59	0.56	0.62	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.02	-0.04	0.08	0.602
Years of Experience	0 - 5 years	-7.77	-20.87	5.34	0.023
	6 - 10 years	-8.73	-23.61	6.16	
	11 - 19 years	3.60	-9.04	16.25	

Note. n= 4531
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 33

Regression Model Associating Junior High Acuity Language Arts Scores with Consistency of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		3.23	0.34	6.11	0.028
Ethnicity	Non-Latino White	7.25	2.23	12.26	0.013
	Latino	4.07	-1.17	9.31	
Economically Disadvantaged		7.47	4.04	10.90	0.000
Prior Year Lang. Arts Score		0.59	0.56	0.62	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		0.55	-0.68	1.78	0.373
Years of Experience	0 - 5 years	-7.46	-19.71	4.80	0.022
	6 - 10 years	-8.24	-22.44	5.96	
	11 - 19 years	3.97	-7.90	15.84	

Note. n= 4531
Note. Consistency of Acuity use is the number of weeks used last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 34

Teacher-Level Variability in Junior High Student Math Scores

Test	Null Model	Full Model (Total Use)	Full Model (Consistency)
AIMS	48.6%	13.5%	13.8%
Acuity predictive tests	33.7%	16.2%	16.5%

Note. *n*(AIMS)= 3814; *n*(Acuity)= 3381

Table 35

Regression Model Associating Junior High AIMS Math Scores with Prevalence of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		-1.59	-3.22	0.04	0.055
Ethnicity	Non-Latino White	2.82	-0.02	5.65	0.001
	Latino	-1.44	-4.36	1.49	
Economically Disadvantaged		4.22	2.24	6.21	0.000
Prior Year Math Score		0.59	0.57	0.61	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		-0.01	-0.03	0.01	0.241
Years of Experience	0 - 5 years	-1.65	-10.72	7.42	0.757
	6 - 10 years	-5.00	-14.37	4.36	
	11 - 19 years	-1.55	-8.38	5.28	

Note. n= 3814
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 36

Regression Model Associating Junior High AIMS Math Scores with Consistency of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		-1.60	-3.22	0.03	0.055
Ethnicity	Non-Latino White	2.84	0.01	5.68	0.000
	Latino	-1.45	-4.37	1.47	
Economically Disadvantaged		4.23	2.24	6.22	0.000
Prior Year Math Score		0.59	0.57	0.61	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		-0.14	-0.64	0.35	0.563
Years of Experience	0 - 5 years	-1.59	-10.79	7.60	0.820
	6 - 10 years	-4.32	-13.69	5.05	
	11 - 19 years	-1.77	-8.75	5.22	

Note. n= 3814
Note. Consistency of Acuity use is the number of weeks used last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 37

Regression Model Associating Junior High Acuity Math Scores with Prevalence of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		-1.37	-4.61	1.88	0.409
Ethnicity	Non-Latino White	1.04	-4.65	6.73	0.095
	Latino	-3.67	-9.49	2.14	
Economically Disadvantaged		5.76	1.82	9.70	0.004
Prior Year Math Score		0.46	0.43	0.50	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		-0.02	-0.06	0.02	0.262
Years of Experience	0 - 5 years	-7.56	-26.47	11.35	0.476
	6 - 10 years	-5.31	-25.95	15.33	
	11 - 19 years	6.36	-8.20	20.92	

Note. n= 3381
Note. Prevalence of Acuity use is the total number of instructional actions performed last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.

Table 38

Regression Model Associating Junior High Acuity Math Scores with Consistency of Acuity Use

		Student-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Gender		-1.37	-4.62	1.88	0.408
Ethnicity	Non-Latino White	1.10	-4.59	6.79	0.089
	Latino	-3.68	-9.50	2.13	
Economically Disadvantaged		5.78	1.83	9.72	0.004
Prior Year Math Score		0.46	0.43	0.50	0.000
		Teacher-Level Variables			
Factor	Category	Mean Difference	95% Confidence Interval		*p*
Prevalence of Acuity Use		-0.40	-1.46	0.66	0.450
Years of Experience	0 - 5 years	-7.21	-26.33	11.91	0.530
	6 - 10 years	-3.93	-24.51	16.65	
	11 - 19 years	6.25	-8.61	21.12	

Note. n= 3381
Note. Consistency of Acuity use is the number of weeks used last year.
Note. Reference category for Gender is female, Ethnicity is Other, Economically Disadvantaged is Yes, and Years of Experience is 20+.