

Galileo Instructional Data System Pilot Project Evaluation

Interim Report

September, 2007



A Measurement Incorporated Company

Galileo Instructional Data System Pilot Project Evaluation

Interim Report

Prepared for:

Massachusetts Department of Education

September, 2007

Prepared by:

MAGI Services

A Measurement Incorporated Company

7-11 South Broadway, Suite 402

White Plains, NY 10601

(914) 682-1969 Fax: (914) 682-1760

Table of Contents

Description of Galileo Instructional Data System Pilot Project.....	I
Evaluation Description and Methodology.....	I
Quality of Program.....	2
Support for Use.....	4
Teacher Participation and Use of Benchmark and Formative Assessments.....	6
School Use of Benchmark Assessments.....	7
Student Interventions.....	9
Summary.....	13
Appendix.....	15

Galileo Instructional Data System Pilot Project Evaluation Interim Report

Description of Galileo Instructional Data System Pilot Project

In 2005, the Massachusetts Department of Education (MADOE) contracted with Assessment Technology, Inc. to supply and implement an instructional data system named Galileo Online. This database system integrates curriculum mapping assessment, reporting and analysis tools, and a standards-based grade book, which allows district and school staff, including teachers, an easy-to-use system for identifying trends in student learning and making improvements in classroom instruction. Key features of the data system are listed below.

- A bank of assessment items aligned with Massachusetts learning standards in Mathematics and English/Language Arts for grades 3-10.
- Tools for constructing district benchmarking assessments and on-demand classroom assessments.
- The ability to analyze, report, and longitudinally track student achievement data.
- Curriculum sequencing and pacing functions.
- Standards-based electronic report card.
- Parent and student web access to student assignments and performance data.

The initial phase of the three-year pilot began in the 2005-06 academic year with 25 schools from 8 districts. During this initial phase, the focus was on the development and use of benchmark assessments. ATI custom designed benchmark assessments based on each district's specification and the state's learning standards. Districts then used the technology to administer and analyze the benchmark assessments. Phase II of the pilot began in 2006-07 and will extend through the third year of the pilot. During this phase, the foci are 1) to engage teachers in classroom formative assessment, and 2) to establish formal systems for student intervention and support.

Evaluation Description and Methodology

The evaluation of the Galileo Instructional Data System Pilot Project (known here on after as Galileo) was conducted by an independent research and evaluation company, MAGI Services, a Measurement Incorporated Company. The two-year evaluation began in the second year of the Galileo pilot (i.e., 2006-2007 academic year) and extends through the third and final year of the pilot project. The evaluation was designed to investigate the quality and implementation of Galileo and its impact on student benchmark outcomes and student intervention and supports.

To this end, the evaluation included participation of district and school staff from all 8 districts and 25 schools in the project. District staff (including Math Support Specialists and district leadership teams) and school administrators/school leadership teams completed an online survey that included items related to the quality of Galileo and use of the system at all levels. Math teachers from each of the 25 participating schools

completed a paper survey with similar items. The response rate for district and school staff was 100% and 50% for teachers from all 25 schools (100% response rate at the school level).

Survey respondents were asked to rate their level of disagreement or agreement for each item on a scale of 1 to 6, with 1 indicating the strongest level of disagreement and 6 indicating the strongest level of agreement. In the findings section of the report, the mean agreement/disagreement score is reported for each of the items. A mean score that fell between a 1.00 and 1.99 is referred to as “strong disagreement”; a score within the 2.00 to 3.99 range is referred to as “disagreement”; a score within the 4.00 to 4.99 range is referred to as “agreement”; and a score that fell between a 5.00 and 6.00 is referred to as “strong agreement.” The mean agreement/disagreement score is reported in tables for each section of the findings. The evaluation chose to be more conservative with the agreement score range and to account for potentially positively-skewed reporting. **Appendix A** includes the percent of respondents who indicated strong agreement to provide additional information about items that leaned more toward high agreement.

This brief presents data related to quality of program, support for use of Galileo, school and teacher participation and use of benchmark and formative assessments, and status of student intervention and remediation services. This data was followed up with an analysis of the student benchmark assessment data. A final summary of the data is provided at the end of the report.

Quality of Program

Quality of Items

- For nearly all of the quality of assessment items listed in **Table 1**, district staff’s responses averaged in the agreement range (4.00 to 4.88). There was one exception, namely that district staff disagreed that there were a sufficient number of items available in the item bank (3.50).
- Teachers’ responses to the quality of benchmark assessment items were within the agreement range (4.10 to 4.60), though they gave slightly lower rates of agreement than the district staff. The one exception was that teachers were more likely to disagree that items were presented in the context of real-world experience and examples (3.97).

Table 1
Mean Scores for Quality of Items

The benchmark assessment items ...	Mean	
	District	Teachers
Are clearly written and easy for students to understand.	4.38	4.14
Are appropriate in terms of difficulty and rigor.	4.38	4.12
Reflect the range of cognitive skills covered by the standards.	4.00	4.14
Are presented in the context of real-world experiences and examples, as appropriate.	4.25	3.97
Are free of bias or sensitivity problems.	4.88	4.60
Have distractors that provide information on student thinking.	4.63	4.39
Seems to be of similar quality as MCAS items.	4.00	4.10
Have significantly improved in quality over the course of the pilot.	4.00	N/A
Are sufficient in the number of available items in the item bank.	3.50	N/A

Sources: online district survey (n=8); teacher survey (n=251)

N/A = not applicable

Alignment with Learning Standards

- As seen in **Table 2**, district staff agreed that assessment items were aligned with Massachusetts learning standards. The average score for each grade level ranged from 4.00 to 4.13.

Table 2
Mean Scores on Alignment with Learning Standards

	Mean
5 th grade	4.00
6 th grade	4.13
7 th grade	4.13
8 th grade	4.13

Source: online district survey (n=8)

Usefulness of Reports

- In all four areas of usefulness of Galileo reports listed in **Table 3**, principals/school leadership teams' responses averaged in the agreement range (4.36 to 4.88).
- Teachers also reported in the agreement range (4.26) for the same items, though their averages were lower than principals/school leadership teams.

Table 3
Mean Scores on Usefulness of Reports

The reports generated by Galileo...	Mean	
	School	Teachers
Contribute to school level instructional planning and decision-making.	4.88	N/A
Contribute to classroom level instruction planning and decision-making.	4.88	4.26
Contribute to student level planning and decision-making.	4.36	4.26
Readily inform instruction and intervention.	4.76	N/A

Sources: online school survey (n=25), teacher survey (n=251)

N/A = not applicable

Ease of Use

- In all areas, district staff, principals/school leadership teams, and teachers agreed that the Galileo system was easy to use (**Table 4**). In some instances, district staff and principal/school leadership teams gave strong agreement ratings. For example, both groups strongly agreed that the data in the system was accurate. Separately, principal/school leadership teams strongly agreed that the reports were clear, concise and easy to understand whereas district staff strongly agreed that staff could extract reports in a timely manner.

Table 4
Mean Scores on Ease of Use

	Mean		
	<i>District</i>	<i>School</i>	<i>Teachers</i>
The reports are clear, concise and easy to understand.	N/A	5.24	4.86
Galileo is well organized and easy to navigate.	4.25	4.72	4.14
The data is accurate.	5.13	5.16	4.63
It is easy to generate reports with the variables for analysis.	4.50	4.52	4.21
It is easy to generate reports in the preferable format.	4.43	4.28	4.09

Sources: online district (n=8), online school survey (n=25), teacher survey (n=251)
N/A = not applicable

Need for Galileo

- As seen in **Table 5**, district staff, principals/school leadership teams strongly agreed (5.20-5.25) that the Galileo assessment system addressed an important need in their school. Teachers agreed (4.27) with this statement as well.

Table 5
Mean Score on Need for Galileo

	Mean		
	<i>District</i>	<i>School</i>	<i>Teachers</i>
Galileo addresses an important need in the school.	5.25	5.20	4.27

Sources: online district (n=8), online school survey (n=25), teacher survey (n=251)

Support for Use

Training and Technology

- District staff agreed that ATI training and materials supported district efforts to provide teachers with the skills needed to use the system overall and to generate reports and analyze data (**Table 6**). To a lesser extent, ATI training supported district efforts to train teachers to develop test items and access items for formative assessment purposes. The average agreement/disagreement score ranged from 4.00 to 4.63.

Table 6
Mean Scores on Training

ATI training supports district efforts to provide teachers the necessary skills to...	Mean
Use the system	4.63
Develop test items	4.00
Access items for formative assessment	4.00
Generate reports	4.50
Analyze and use data to inform instruction	4.38

Source: online district survey (n=8)

- As seen in **Table 7**, principals/school leadership teams agreed that the technology resources at the school were in satisfactory condition. The mean agreement/disagreement score ranged from 4.16 to 4.88.

Table 7
Mean Scores on School Technology Resources
Mean Scores on Training

School Technology Resources	Mean
Financial resources are allocated to support installation and ongoing updates of technology and networking infrastructure.	4.36
Human resources are allocated to support installation and ongoing updates of technology and networking infrastructure.	4.04
Trained staff are available to provide site-based technical and troubleshooting support for technology use.	4.24
Professional development is offered to staff on how to use general computer technology.	4.16
Teachers have easy access to updated, internet accessible computers.	4.88
Scanners are in good order and are readily available to staff for scanning benchmark assessments.	4.25

Source: online school survey (n=25)

Time Provided to Use Galileo Effectively

- Principals/school leadership teams had mixed agreement over the amount of time teachers had to use Galileo and the assessment data (**Table 8**). They agreed that teachers had sufficient time to administer the benchmark assessments (4.40), use the data to reflect on instruction (4.29), and to work with colleagues to exchange successful instructional practices and means of improvement (4.17).
- On the other hand, principals/school leadership teams disagreed with statements regarding time to conduct data-based, collaborative discussions on student data (3.96) and time to meet with each other to design formative assessments (3.21).

Table 8
Mean Scores on Time Provided to Use Galileo Effectively
As reported by the School Survey

There is sufficient time for teachers to...	Mean
Administer the benchmark assessments.	4.40
Conduct data-based, collaborative discussions focusing on student data and progress.	3.96
Use benchmark assessment data to collaboratively reflect on instruction.	4.29
Work with colleagues to exchange successful instructional strategies and to discuss means of improving and enhancing teacher practices.	4.17
Meet with each other to design formative assessments.	3.21

Source: online school survey (n=25)

- Teachers were in more disagreement about having sufficient time to use Galileo effectively compared to principals/school leadership teams (**Table 9**). The only area where teachers felt that they had sufficient time was in administering the benchmark assessments (4.20).
- Otherwise, most teachers felt that they did not have sufficient time to develop and administer formative assessments (3.36) or to review the data from benchmark or formative assessments (3.71). Furthermore, they did not have sufficient time to plan instructional activities around areas of student weaknesses' (3.39) or to re-address weaknesses with students individually or within a group (3.12 and 3.37, respectively).

- Similar to principals/school leadership teams, teachers reported that they did not have sufficient time to collaborate with other teachers to analyze assessment data from Galileo.

Table 9
Mean Scores on Time Provided to Use Galileo Effectively
As reported by the Teacher Survey

There is sufficient time for teachers to...	Mean
Administer the benchmark assessments.	4.20
Develop and administer formative assessments.	3.36
Review the data from the assessments.	3.71
Plan instructional activities to address areas of student weaknesses.	3.39
Re-address areas of weaknesses with a student.	3.12
Re-address areas of weaknesses with a group of students.	3.37
Use differentiated instruction to meet learner’s needs based on the data.	3.63
Collaborate with other teachers to analyze assessment data from Galileo.	3.35

Source: Teacher Survey (n=251)

Teacher Participation and Use of Benchmark and Formative Assessments

Teacher Participation in Development and Analyses

- As seen in **Table 10**, teachers’ responses to participation in the development and analyses of benchmark and formative assessments fell mostly in the disagreement range. In one area — development of school-wide or grade-level formative assessments — teachers strongly disagreed. The average agreement/disagreement scores ranged from 2.95 to 3.80.

Table 10
Mean Scores on Teacher Participation in Development and Analyses

Teachers participate in...	Mean
The identification of target standards for benchmark assessment and/or review of benchmark assessments.	3.80
The developments of school-wide or grade-level formative assessments.	2.95
Formal meetings to collaboratively analyze benchmark assessment data reports.	3.97
Formal meetings to collaboratively analyze school-wide or grade level formative assessment data reports.	3.40
Development of formative assessments for my classroom.	3.13
Development and/or review of my own classroom formative assessment data reports.	3.38
Collaborative review of student work on open response items to analyze student thinking.	3.46

Source: Teacher Survey (n=251)

Teacher Use of Benchmark and Formative Assessments

- Both principals/school leadership teams and teachers agreed that teachers used benchmark assessments to evaluate the progress of students and to identify struggling students (**Table 11**).
- Teachers also agreed that benchmark assessments were used for the purposes of adjusting curriculum in areas where students' encountered problems and for differentiating instruction based on student needs.

Table 11
Mean Scores on Teachers' Use of Benchmark Assessments
Mean Scores for Quality of Items

Teachers use benchmark assessments to...	Mean	
	School	Teacher
Adjust curriculum in areas where students encountered problems.	4.21	4.08
Differentiate instruction based on student needs.	3.54	4.08
Evaluate the progress of students.	4.42	4.44
Place students in instructional groups.	3.67	3.56
Identify struggling students.	4.58	4.42

Sources: Online School Survey (n=25), Teacher Survey (n=251)

- Teachers and principal/school leadership teams were less likely to agree with statements related to the use of formative assessments (**Table 12**). Teachers did, however, use formative assessments to evaluate the progress of students and to identify struggling students.

Table 12
Mean Scores on Teachers' Use of Formative Assessments

Teachers use formative assessments to...	Mean	
	School	Teacher
Adjust curriculum in areas where students encountered problems.	3.91	3.83
Differentiate my instruction based on student needs.	3.70	3.85
Evaluate the progress of students.	4.13	4.04
Place students in instructional groups.	3.65	3.44
Identify struggling students.	4.35	4.01

Sources: Online School Survey (n=25), Teacher Survey (n=251)

School Use of Benchmark Assessments

- As seen in **Table 13**, principals/school leadership teams agreed that benchmark assessment data was used to monitor the performance and progress of students, classrooms, and grade levels, but less so in the form of disaggregated data.
- Benchmark assessments were also used more for purposes of setting goals and developing plans for long term improvement, more so than for immediate goal setting.

- In the area of evaluation, principals/school leadership teams agreed that benchmark assessment data was used to identify student problems, verify causes and generate and monitor solutions. The data were also used to evaluate the math curriculum and modify curriculum and pacing guides.
- In the area of programming, the assessment data was more likely to be used to make scheduling decisions rather than to inform budget decisions or to make staffing decisions.
- Benchmark assessment data was used to determine priorities for professional development, both for the whole school and for individual teachers.
- In the area of student interventions, principals/school leadership teams agreed that assessment data was used to identify students in need of math supplemental intervention and remediation. The data was less likely to be used to monitor students' progress in supplemental intervention and remediation, which included determining when services were no longer needed and developing individualized student profiles.

Table 13
Mean Scores on Schools' Use of Benchmark Assessments

The school uses benchmark assessments to...	Mean
Monitor	
Monitor the performance and progress of each student, classroom, and grade level throughout the year.	4.04
Monitor the performance and progress of disaggregated groups of students throughout the school year.	3.63
Set Goals and Objectives	
Set demanding, objectively measurable goals for immediate improvement	3.79
Develop plans and timeframes for reaching immediate goals.	3.88
Set demanding, objectively measurable goals for long-term improvement.	4.08
Develop plans and timeframes for reaching long-term goals.	4.08
Evaluate	
Identify student learning problems, verify causes and generate and monitor solutions.	4.04
Evaluate the math curriculum and modify curriculum and pacing guide calendars as needed.	4.04
Programming	
Inform budget decisions.	3.96
Make staffing decisions.	3.71
Make scheduling decisions about the use of teacher and student time.	4.33
Professional Development	
Determine priorities for professional development for the school as a whole.	4.54
Determine priorities for professional development for individual teachers.	4.13
Student Interventions	
Identify students in need of math supplemental intervention and remediation.	4.50
Monitor students' progress in math supplemental intervention and remediation.	3.88
Determining when math supplemental intervention and remediation services are no longer needed.	3.70
Developing individualized student profiles to inform supplemental intervention and remediation.	3.54

Source: Online School Survey (n=25)

Student Interventions

- As seen in **Table 14**, schools agreed that there were guidelines for placing students in and out of supplemental intervention and remediation services.
- Classroom grades had the highest priority in terms of identifying students for interventions. Grades were followed by teacher recommendations, standardized achievement test scores and benchmark assessment data.
- Less likely to be used for identifying students was formative assessment data, other staff recommendations, requests from parents, and student portfolios.

Table 14
Mean Scores on Identification of Students for Intervention and Remediation Services

	Mean
There are systematic guidelines for placing students both into and out of supplemental intervention and remediation services.	4.33
The following are used to identify students for interventions:	
a. benchmark assessment data	4.57
b. formative assessment data	3.96
c. standardized achievement test scores	4.70
d. classroom-based grades	5.08
e. student portfolios	2.50
f. teacher recommendations	4.79
g. other staff recommendations	3.83
h. requests from parents	3.61

Source: Online School Survey (n=25)

- As seen in **Table 15**, not all students who were in need of additional support received it. In 13 out of 24 schools, more students needed supplemental intervention and remedial services than received them. Conversely, in 9 schools, the same percentage of students who needed services received them.

Table 15
Percent of Students in Need of and Receiving Supplemental Intervention and Remedial Services

	Number of Schools indicating Students in Need	Number of Schools indicating Student Participation
21-30% of students	4	1
31-50% of students	8	2
51-75% of students	7	3
76% + of students	3	3

Source: Online School Survey (n=25)

Looking at **Table 16**...

- Regular classroom teachers provided supplemental intervention services and remediation to students in the large majority of schools (92%).
- In just over half of schools, school or district-based specialists and teachers' aides also provided supplemental instruction.

Table 16
Type of Staff Available to Provide Supplemental Intervention Services and Remediation to Students

	Percent of Schools
Regular classroom teachers	92% (23)
School-or district-based specialist	52% (13)
Teachers' aide	52% (13)
External professional provider	32% (8)
Volunteers	20% (5)

Source: Online School Survey (n=25)

- As seen in **Table 17**, the top three types of supplemental instruction for students in need included after school programs (80%), pull-out during the day (76%), and summer school (64%).

Table 17
Time when Supplemental Intervention and Remediation Services are Available

	Percent of Schools
After school	80% (20)
Pull-out during the day	76% (19)
Summer school	64% (16)
Before school	24% (6)
Saturday	12% (3)
Intersession program	8% (2)

Source: Online School Survey (n=25)

- Over half (61%) of schools reported that students averaged between 2 to 3 hours of participation in supplemental and remediation services (**Table 18**).

Table 18
Hours per Week of Student Participation in Supplemental Intervention and Remediation Services

	Percent of Schools
1 hour per week	4% (1)
2-3 hours per week	61% (14)
4-5 hours per week	26% (6)
More than 5 hours per week	9% (2)

Source: Online School Survey (n=25)

Student Benchmark Analyses

Students in grades 5 through 8 were administered district-based, tailor made benchmark assessments quarterly throughout the school year. Of the 25 schools, only 8 administered the fourth quarter assessment, therefore the analyses focused on the third quarter assessment as the last assessment of the school year.

The analyses of the student benchmark assessments included comparisons of three groups. The first group included teachers who ranked high and low on use of benchmark assessments in the classroom. High and low groups were defined by the first and fourth quartile scores of the implementation scale. Implementation of benchmark assessment data in the classroom was deconstructed into the five statements outlined in **Table 11** of this report. ANCOVA analyses were performed with the first benchmark assessment score as a covariate.

As seen in **Table 19**, the estimated marginal mean for the third quarter benchmark test in the high implementing group was 1156.95 whereas the mean for the low group was 1140.27; the mean difference was 16.68 points. The difference between the means was statistically significant. In other words, students performed higher on the third quarter benchmark test in classrooms where teachers made greater use of benchmark assessment data to drive their instruction in the classrooms as compared to teachers who used the data less throughout the school year.

Table 19
Estimated Marginal Means (and Standard Error) of the High and Low Implementing Teacher Groups

	N	Mean	Standard Error
Low Implementing Teacher Group	45	1140.27	6.22
High Implementing Teacher Group	54	1156.95	5.67

The second comparison was based on the percent of students who were receiving intervention and remediation services. The high group was defined as 51% or more of students and the low group was defined as 30% or less of students in intervention and remediation. ANCOVA analyses were performed with the first benchmark assessment score as a covariate.

As seen in **Table 20**, the estimated marginal mean for the third quarter benchmark test in the high student group was 1167.98 whereas the mean for the low group was 1148.17; the mean difference was 19.81 points. The difference between the means was statistically significant. In other words, students performed higher on the third quarter benchmark test in schools that had a higher percent of students receiving intervention and remediation services.

Table 20
Estimated Marginal Means (and Standard Error) of the High and Low Student Groups

	N	Mean	Standard Error
Low Student Group	13	1148.17	4.67
High Student Group	6	1167.98	7.01

The third group comparison was based on schools where the percent of students receiving intervention and remediation was relative to the percent in need. The high group included schools where the same percent or

more of students who needed services also received them. The low group included schools where a lower percent of students received services than needed the services. ANCOVA analyses were performed with the first benchmark assessment score as a covariate.

As seen in **Table 21**, the estimated marginal mean for the third quarter benchmark test in the high student group was 1165.99 whereas the mean for the low student group was 1152.78. The difference between the two groups was 13.21 points; however this difference was not statistically significant. The finding does point to a trend of higher benchmark scores in schools where the same percent or more of students who needed services also received them.

Table 21
Estimated Marginal Means (and Standard Error)
of the High and Low Student Groups

	N	Mean	Standard Error
Low Student Group	13	1152.78	4.64
High Student Group	9	1165.99	5.58

Summary

Across the board, district staff, principals/school leadership teams, and teachers agreed with the quality, alignment, usefulness, and ease of use of the Galileo assessment system. Also important, district staff and principals/school leadership teams strongly agreed and teachers agreed that Galileo addressed an important need in the school. Taken together, this data suggests that there is buy-in from the districts and from schools for use of the program.

Like any other pilot project, continuous improvement of the system is important for reinforcing buy-in and continued use of the benchmark assessment system. While most constituents were generally satisfied with the system, there were areas where the data showed a need for some improvement. One area that district staff pointed out as a needed improvement was in the number of benchmark assessment items in the item bank.

Alignment of the benchmark assessments with the Massachusetts learning standards is certainly critical to the future use of Galileo. Though district staff agreed that the items were aligned with the standards, the mean scores hovered in the low average range (4.00 to 4.13). One would expect that a critical component of the program would have higher levels of agreement in order to be successful in the long term. Therefore, we recommend that ATI assure district and school staff that there is strong alignment between the items and the standards through their reports.

The evaluation also looked at areas of support for use of Galileo, including training, technology, and provision of time. On the positive side, district staff agreed that ATI training supported their efforts to provide teachers with the necessary skills to use the system, generate reports, and analyze the data and, to a lesser extent, develop test items and access items for the formative assessment. Furthermore, principals/school leadership teams agreed that technology resources were satisfactory.

On the other hand, apart from adequate time to administer the benchmark assessments, teachers felt that they were not provided with the time needed to analyze and use the assessment data. Specifically, teachers did not have time to review the data or to use the data for instructional planning, to address weaknesses with students, or to collaborate with fellow colleagues to analyze the data. Principals/school leadership teams also agreed that teachers did not have the time needed to conduct data-based collaborative discussions with others and meet with each other to design formative assessments.

Perhaps as a result of insufficient time, teacher participation in the development and analyses of the data was low. Teachers were unlikely to participate in the identification of target standards for benchmark assessments, development of school-wide or grade-level formative assessments, and collaborative review of student work on open response items. Teachers were also less likely to participate in formal meetings to collaboratively analyze benchmark assessment reports and school-wide or grade level formative assessment data reports. Finally, teachers were less likely to develop formative assessments for their classroom and to review formative assessment data reports.

Regardless of the limited time and participation in the development and analyses of assessment data, teachers and principals/school leadership teams agreed that benchmark assessment data was used in the classroom to make adjustments in the curriculum, differentiate instruction, evaluate student progress, and identify struggling students. Teachers were less likely to use formative assessments for the same purposes, with the exception of using them to evaluate the progress of students and identify struggling students. The limited use of formative assessment data is not surprising in light of the data that showed teachers were less likely to develop and review formative assessments.

If teachers did not participate in development and analyses of assessments, then who did and how did this data get to teachers? It is most likely that district and school administrators and leadership teams had primary responsibility over the development and analyses of the data. A summary of the analyses should have been provided to teachers to be used in the classroom to evaluate student progress and identify struggling students.

We would encourage greater teacher participation in development and analyses of assessments. Galileo is designed to allow user-friendly data access at all levels, meaning that teachers can use these tools to engage in the informed reflection necessary to improve classroom practice, if given the time and opportunity to do so. With this new technology that allows efficient organization and access to student data, all district and school staff can and should be involved in data analysis. Research suggests that greater involvement in data analysis by teachers will result in more use of data in the classroom. In turn, as this study showed, more use of data in the classroom can and did result in higher scores.

Greater teacher participation can be achieved in several ways.¹ One suggestion is to identify people in the school who want to use and who see the value of the data system and train them to be proficient in using and analyzing the data. Teachers are more likely to be inspired by a fellow colleague than an outside consultant or district personnel. Teachers also need good training on how to ask the right questions and how to answer such questions through mining the data and not just on how to use the online database system. Another critical component is to provide the time to look at the data, identify areas in need of improvement and most importantly, use the data thoughtfully to improve instruction. Teachers should have opportunities to share strategies; for instance, if one classroom did particularly well on a standard, the teacher in that class should be encouraged to share her instructional approaches with other teachers. Use of the data should be driven by inquiry and not by fear and this can be achieved by providing opportunities for teachers to open up about what is working or not working on their classrooms.

As for the student intervention and remediation data, the most striking finding was that more students were in need of intervention or remediation than those who actually received these services. This may be a reflection of low participation in after-school programs, which was the primary type of intervention or remediation in these schools. Indeed, the data showed that in most schools, students participated in 2-3 hours of intervention or remediation weekly. The low percentages may also be a reflection of limited resources to enroll students into these programs.

The student benchmark data analyses revealed two significant findings. One, in classrooms where teachers reported higher use of benchmark assessment data in their instruction, students also performed higher on the third quarter assessment as compared to teachers who reported lower use of the assessment data. This finding suggests that a focus on using data to drive instruction can result in improved student achievement.

Two, in schools where a higher percentage of students (e.g., 51% or more) were enrolled in intervention and remediation services, students also performed higher on the third quarter benchmark assessment as compared to schools with lower percentages of students enrolled in these programs. This finding underscores the value of intervention and remediation; therefore, more work is certainly needed to ensure students that need extra support receive it. If programs, particularly before- and after-school programs, are operating on a drop-in basis, attendance is likely to be sporadic and attrition is likely to be high. We encourage schools to mandate attendance and continuous enrollment into these programs. If transportation is an issue, busing opportunities could be made available. School staff should also communicate the value of these programs to parents in order to encourage more enrollment.

¹ Suggestions for teacher participation were derived from *Using data: The math's not the hard part*. Learning Points Associates, 2006

Appendix

Quality of Items

The benchmark assessment items...	% of strong agreement	
	<i>District</i>	<i>Teachers</i>
Are clearly written and easy for students to understand.	38% (3)	41% (102)
Are appropriate in terms of difficulty and rigor.	75% (6)	40% (100)
Reflect the range of cognitive skills covered by the standards.	50% (4)	43% (108)
Are presented in the context of real-world experiences and examples, as appropriate.	50% (4)	36% (88)
Are free of bias or sensitivity problems.	75% (6)	63% (154)
Have distractors that provide information on student thinking.	63% (5)	51% (122)
Seems to be of similar quality as MCAS items.	50% (4)	47% (116)
Have significantly improved in quality over the course of the pilot.	50% (4)	N/A
Are sufficient in the number of available items in the item bank.	38% (3)	N/A

N/A = not applicable

Alignment with Learning Standards

	% of strong agreement
5 th grade	60% (3)
6 th grade	50% (4)
7 th grade	50% (4)
8 th grade	50% (4)

Usefulness of Reports

The reports generated by Galileo	% of strong agreement	
	<i>School</i>	<i>Teachers</i>
Contribute to school level instructional planning and decision-making.	72% (18)	N/A
Contribute to classroom level instruction planning and decision-making.	72% (18)	46% (113)
Contribute to student level planning and decision-making.	52% (13)	48% (117)
Readily inform instruction and intervention.	64% (16)	N/A

N/A = not applicable

Ease of Use

	% of strong agreement		
	<i>District</i>	<i>School</i>	<i>Teachers</i>
The reports are clear, concise and easy to understand.	N/A	80% (20)	69% (170)
Galileo is well organized and easy to navigate.	38% (3)	60% (15)	43% (100)
The data is accurate.	75% (6)	80% (20)	62% (144)
It is easy to generate reports with the variables for analysis.	50% (4)	64% (16)	41% (91)
It is easy to generate reports in the preferable format.	57% (4)	52% (13)	40% (90)
Staff can extract reports in a timely manner.	88% (7)	71% (17)	55% (127)

N/A = not applicable

Need for Galileo

	% of strong agreement		
	<i>District</i>	<i>School</i>	<i>Teachers</i>
Galileo addresses an important need in the school	88% (7)	84% (21)	47% (111)

Training and Technology

ATI training supports district efforts to provide teachers the necessary skills to...	% of strong agreement
Use the system	63% (5)
Develop test items	25% (2)
Access items for formative assessment	50% (4)
Generate reports	63% (5)
Analyze and use data to inform instruction	50% (4)

School Technology Resources	% of strong agreement
Financial resources are allocated to support installation and ongoing updates of technology and networking infrastructure.	52% (13)
Human resources are allocated to support installation and ongoing updates of technology and networking infrastructure.	40% (10)
Trained staff are available to provide site-based technical and troubleshooting support for technology use.	56% (14)
Professional development is offered to staff on how to use general computer technology.	48% (12)
Teachers have easy access to updated, internet accessible computers.	84% (21)
Scanners are in good order and are readily available to staff for scanning benchmark assessments.	42% (10)

Time Provided to Use Galileo Effectively

There is sufficient time for teachers to...	% of strong agreement
Administer the benchmark assessments.	56% (14)
Conduct data-based, collaborative discussions focusing on student data and progress.	42% (10)
Use benchmark assessment data to collaboratively reflect on instruction.	46% (11)
Work with colleagues to exchange successful instructional strategies and to discuss means of improving and enhancing teacher practices.	42% (10)
Meet with each other to design formative assessments.	21% (5)

There is sufficient time for teachers to...	% of strong agreement
Administer the benchmark assessments.	51% (126)
Develop and administer formative assessments.	21% (52)
Review the data from the assessments.	31% (77)
Plan instructional activities to address areas of student weaknesses.	23% (56)
Re-address areas of weaknesses with a student.	18% (43)
Re-address areas of weaknesses with a group of students.	21% (51)
Use differentiated instruction to meet learner's needs based on the data.	27% (67)
Collaborate with other teachers to analyze assessment data from Galileo.	24% (60)

Teacher Participation in Development and Analyses of Assessments

Teachers participate in...	% of strong agreement
Administer the benchmark assessments.	51% (126)
Develop and administer formative assessments.	21% (52)
Review the data from the assessments.	31% (77)
Plan instructional activities to address areas of student weaknesses.	23% (56)
Re-address areas of weaknesses with a student.	18% (43)
Re-address areas of weaknesses with a group of students.	21% (51)

Teachers' Use of Benchmark Assessments

Teachers use benchmark assessments to...	% of strong agreement	
	School	Teachers
Adjust curriculum in areas where students encountered problems.	46% (11)	47% (117)
Differentiate instruction based on student needs.	63% (15)	44% (110)
Evaluate the progress of students.	54% (13)	57% (141)
Place students in instructional groups.	29% (7)	37% (92)
Identify struggling students.	58% (14)	60% (150)

Teachers' Use of Formative Assessments

Teachers use formative assessments to...	% of strong agreement	
	<i>School</i>	<i>Teachers</i>
Adjust curriculum in areas where students encountered problems.	26% (6)	43% (107)
Differentiate instruction based on student needs.	13% (3)	42% (103)
Evaluate the progress of students.	39% (9)	51% (127)
Place students in instructional groups.	30% (7)	35% (87)
Identify struggling students.	48% (11)	52% (128)

Schools' Use of Benchmark Assessments

The school uses benchmark assessments to...	% of strong agreement
Monitor	
Monitor the performance and progress of each student, classroom, and grade level throughout the year.	38% (9)
Monitor the performance and progress of disaggregated groups of students throughout the school year.	33% (8)
Set Goals and Objectives	
Set demanding, objectively measurable goals for immediate improvement	25% (6)
Develop plans and timeframes for reaching immediate goals.	29% (7)
Set demanding, objectively measurable goals for long-term improvement.	42% (10)
Develop plans and timeframes for reaching long-term goals.	38% (9)
Evaluate	
Identify student learning problems, verify causes and generate and monitor solutions.	33% (8)
Evaluate the math curriculum and modify curriculum and pacing guide calendars as needed.	38% (9)
Programming	
Inform budget decisions.	46% (11)
Make staffing decisions.	33% (8)
Make scheduling decisions about the use of teacher and student time.	50% (12)
Professional Development	
Determine priorities for professional development for the school as a whole.	71% (17)
Determine priorities for professional development for individual teachers.	54% (13)
Student Interventions	
Identify students in need of math supplemental intervention and remediation.	63% (15)
Monitor students' progress in math supplemental intervention and remediation.	42% (10)
Determining when math supplemental intervention and remediation services are no longer needed.	22% (5)
Developing individualized student profiles to inform supplemental intervention and remediation.	33% (8)

