DOCUMENT RESUME

ED 452 221 TM 032 506

AUTHOR Cromey, Allison; Hanson, Matthew
TITLE An Exploratory Analysis of School-Based Student Assessment Systems.
INSTITUTION North Central Regional Educational Lab., Oak Brook, IL.
PUB DATE 2000-02-00
NOTE 59p.
AVAILABLE FROM North Central Regional Educational Laboratory (NCREL), 1900 Spring Road, Suite 300, Oak Brook, IL 60523. Tel: 630-571-4700; Fax: 630-571-4716.
PUB TYPE Reports - Evaluative (142) -- Tests/Questionnaires (160)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Decision Making; *Educational Assessment; Elementary Secondary Education; *Student Evaluation; *Systems Development; Test Construction; *Test Use

ABSTRACT

This study examined the general nature of school-based student assessment in a small sample (five elementary and three middle schools and one high school). It relies on descriptive analyses and the presentation of data obtained during site visits. The schools have been separated into those with well-developed student assessment systems and those that do not have such systems. Schools in the well-developed group were found to be very similar in the general strategies they used to align their standards and curriculum to the multiple student assessments they used. They were also similar in the purposes for which they used student assessment results for instructional decision making and to validate all assessment tools. These schools were serious about committing time and resources to developing teachers' and administrators' capacity to reflect on and use assessment data, and they worked to maximize the efficiency of their assessment systems while recognizing that they would never be finished designing curricula and assessment systems. Schools in the less developed group had made less progress in aligning the various elements of the assessment system and had not articulated a clear purpose for the assessments they used. Appendixes contain the protocols for the teacher and administrator visit interviews, the school findings summaries, and a teaching for mastery teacher outline. (Contains 2 tables, 3 figures, and 23 references.) (SLD)

An Exploratory Analysis of School-Based Student Assessment Systems

February 2000

By
Allison Cromey and Matthew Hanson,
North Central Regional Educational Laboratory

Contents

Appendices

Tables

Figures

Introduction

Achievement testing has long been a major fixture of the American educational landscape. Tests are used to monitor the quality of school systems, evaluate education policies and programs, make important instructional decisions about students, and leverage major educational change (Baker & Linn, 1997; Linn, 1986). In recent years, student testing—particularly from the point of view of the teacher—has become increasingly complex as classrooms have become inundated by growing numbers of mandated tests, new forms of student assessment, and a host of educational reform initiatives that demand accountability and are linked to testing (Cizek, 1995).

Educators, reformers, and researchers have begun to recognize some of the challenges schools face using multiple student assessments. For example, coordinating assessment activities and aligning them with a common "vision" for student achievement can be very demanding work for schools, particularly when, as research has consistently shown, few teachers and administrators have received formal training in assessment (Resnick & Resnick, 1989; Schafer & Lissitz, 1987; Sokoloff, 1987; Stiggins, 1995; Stiggins & Conklin, 1992; Wise, Lukin, & Roos, 1991). The national organizations, the research community, and state and local education agencies have attempted to provide guidance to schools regarding the use of student assessments. Nevertheless, it not well understood how—or how well—schools are using these supports.

This report addresses some of these issues by summarizing findings from an exploratory study of the school-based assessment practices in a sample of elementary, middle, and secondary schools. The purposes of the study were twofold: (1) to add to the growing base of knowledge about how schools use student assessment data obtained from multiple sources to inform important decisions about programs, instruction, and individual students; and (2) to identify and describe the factors and conditions that make schools' use of the student assessment data more probable and valuable. Specifically, four variables were examined in this study:

1. The *indicators of performance* used by schools to assess student achievement in core subject areas such as mathematics, reading, and writing
2. The *users* of assessment information
3. The *purposes* for which student assessment data are used
4. The *processes* that support and enhance school staff competency in using student assessments and interpreting their results

These variables can be used in the future as the basis for a more comprehensive evaluation of school-based student assessment practices in districts and states. In the near term, the findings summarized in this report are intended to help state, regional, and local education agencies "take stock" of their current capacity to assist schools in developing, implementing, and overseeing student assessment systems; and their ability to help decision makers maximally use the information garnered from these systems.

First, we will summarize recent trends in state testing and the accountability uses of assessment results to provide a broader context for understanding the environment in which school-based student assessment systems are often situated.

Background/Context

Policymakers, educators, and the general public are looking once again to student assessment as a catalyst for education reform because of its reputed ability to leverage instructional improvement and to hold school systems accountable for their results. Recent mandates for improved student outcomes have shifted focus away from assessments that test students' "minimal" skills to those that emphasize high standards in skill areas that are considered necessary for success in today's technologically advanced society. Specifically, today's educational reforms are often linked to state frameworks that broadly specify the content and knowledge expectations for all students. In addition, state frameworks also serve as guidance for local curriculum development, textbook selection, teacher professional development, and evaluation of schools (Baker & Linn, 1997).

The standards movement, which has origins in national education initiatives, such as *America 2000* and its predecessor, *Goals 2000,*[1] can be attributed to research on how children learn, the effects of minimum competency testing on instruction, and on the public's demand to hold schools accountable for student achievement (Darling-Hammond & Wise, 1985; Glaser & Silver, 1994; Herman, Dreyfus, & Golan, 1990; Herman & Golan, 1991; Madaus, 1988; Shepard, 1989). In response, states have, one by one, made student assessment the centerpiece of their school reform and improvement efforts. Today, 49 states have implemented a statewide exam—46 of these measure student achievement relative to specified content and performance standards in core subjects such as reading, math, and science (Council of Chief State School Officers, 1998). Increasingly, state assessments are being linked to policies that hold school systems, teachers and administrators, and the students themselves accountable for students' performance. Students' scores on state assessments are, for example, reported in the newspaper and other media as a matter of public accounting; are used as a basis for determining rewards and sanctions for schools and their staffs; and are used in making decisions about whether to graduate students or promote them to the next grade.

The increased use of state assessments as a basis for school improvement and accountability has placed new demands on schools that already use their own assessments to measure what their students know and can do. These school exams include (1) *classroom* assessments such as teacher-made quizzes and tests and informal observations of students' learning; (2) *school assessments* such as student writing portfolios or other projects that are used to unify the curriculum and assess student proficiencies in multiple subject areas; and (3) *district assessments* such as commercially purchased standardized achievement tests that can be administered under uniform conditions in many schools simultaneously. Thus, teachers and administrators can easily be inundated with information about the performance of their students.

To prevent this, schools must implement assessment systems that are not only aligned and integrated with local curricula, instructional practices, and professional development strategies, but also contribute to the goal of increasing student achievement based on rigorous content standards. This is complex, demanding work that can take several years. Similarly, managing, synthesizing, interpreting, and using student assessment data obtained from a multifaceted

[1] See also National Council on Education Standards and Testing, 1992; National Education Goals Panel, 1991; and U.S. Department of Education, 1991, 1994.

assessment system can be a daunting task for educators, particularly when (a) the assessment system lacks coherence; (b) school staff do not have training or experience in student assessment; and (c) the time, attention, and energies of teachers and administrators are stretched to personal limits.

Method

Participants

A total of nine public schools—five elementary, three middle, and one high school—were selected for the study based on nominations made by Michigan State Board of Education staff. Specifically, they were asked to identify schools they thought represented a range of experiences relating to:

- Using assessments that are aligned with district and/or state standards.
- Using assessments as part of schoolwide improvement efforts.
- Having current or prior teacher involvement in assessment development or related activities.
- Providing professional development support for teachers using student assessments and their results.
- Using assessment data as the basis for making instructional, program, and policy decision making.

Additional measures were taken to maximize the representation of schools in this study. Selected schools were located in different regions of the state and were diverse in terms of the size of their student enrollments and key student characteristics (e.g., SES, ethnicity). Table 1 summarizes some of the key characteristics of schools that were visited by NCREL researchers.

A total of 46 interviews were conducted in the nine sites. The interviewees included regular, bilingual, and Title I teachers; building principals; district curriculum and assessment coordinators; assistant superintendents; and, in one case, a district superintendent. Teachers were selected by their district's curriculum director or principal based on their prior experience with using student assessments and data. For example, some teachers were selected for interview because they had been involved in developing assessments for their school or district. Others were members of their school's improvement team, participated in local efforts to align the curriculum with standards, or served on district or state committees for content standards and benchmarks. In addition, at each school, some teachers were purposely chosen because they had little or no experience in any of these areas.

NCREL is aware that it is not possible to "see the whole picture" without interviewing more educators in more schools. Moreover, we acknowledge the fact that schools and individuals were not selected randomly for the study. Quite the contrary, they were, in addition to the criteria shown above, identified by SEA staff based on their reputations, personal contacts with the agency, and/or

general willingness to participate in studies like ours—criteria that may reflect personal biases. Nonetheless, we do maintain that the sample is adequate to explore issues relating to student assessment and is consistent with the formative and exploratory nature of this study.

Procedures

Two NCREL researchers visited each of the nine schools. Semistructured interview protocols (see Appendices A and B) were used to guide discussions with teachers, principals, and district-level staff. These interviews typically took one day to complete. In addition, school and district staff were asked to provide any documents they thought would help researchers better understand the student assessments they develop and/or use locally, as well as their efforts to set standards, align curriculum, or both. All study participants were assured that their identities and those of their schools would be kept confidential. Therefore, all direct quotations and documents used in this report are not traceable to their source by name.

Table 1. Characteristics of Participating Schools

	Level, Location[a]	% F/R Lunch[b]	% Minority (non-caucasian)	Teacher/ Pupil Ratio	$/Pupil[c]	Avg Tchr Salary[d]
1999 State Average		31		22	4,400	47,700
School A	E, Rur	34	3	17	4,600	41,600
School B	E, Sub	9	6	22	4,100	56,600
School C	E, Urb	90	23	30	4,300	56,700
School D	E, Sub	17	30	23	4,900	51,400
School E	E, Rur	9	5	24	4,500	55,500
School F	MS, Rur	28	3	20	4,500	45,500
School G	MS, Rur	26	6	22	4,900	49,900
School H	MS, Urb	50	47	18	4,700	46,200
School I	HS, Sub	2	6	20	4,900	47,300

[a] E = Elementary, MS = Middle School, HS = High School; Rur = Rural, Urb = Urban, Sub = Suburban
[b] Free/Reduced Lunch (Rounded)
[c] Expenditure per pupil (Rounded)
[d] Average Teacher Salary (Rounded)

Data Analysis

Researchers' notes, direct quotations obtained from informants, and relevant documents collected from schools were analyzed for content and theme. One-page, school-level summaries of these notes are included in Appendix C.

Findings

The findings presented in this report are the result of cross-site analyses of interview and document review data obtained from nine schools in Michigan. The results of these analyses are presented according to key characteristics that, *together*, distinguish schools that have made significant progress in implementing school-based student assessment systems from those that are, to date, not as far along. For the sake of parsimony, we culled schools into two groups based on these characteristics and identified them as having "well-developed" (N=4) or "less-developed" (N=5) assessment systems.

Specifically, schools with well-developed assessment systems distinguish themselves from schools with less-developed assessment systems because:

1. They align local curriculum, standards, and assessments to state content standards.
2. They analyze assessment results to monitor student progress.
3. They use state assessment results to check the validity of local assessment systems.
4. They use assessment results to evaluate the efficacy of local curriculum and instructional practices.
5. They limit the number of student assessments used to those that are purposeful and can be aligned with local curriculum and state standards.
6. They allocate time for teachers to collaborate, reflect, and make data-based decisions—individually or in teams—based upon student assessment data and their instructional implications.

The majority of the findings section will focus on the schools with well-developed assessment systems in an attempt to describe not only what they are doing to develop these systems, but also how they are managing this challenging work. The schools with less-developed assessment systems were either not as advanced as the well-developed schools or were grappling with the challenges associated with implementing their own school-based assessment systems. The variability in the quality of the assessment systems at these schools and in the issues they faced was substantial. Detailing all of these differences is beyond the scope of this paper. Instead, a brief description within the section for each major finding will contrast how the characteristic was manifested in these less-developed schools and, where applicable, what important barriers related to the issue. Examples of school activities and quotations from informants will be used to illustrate findings throughout this report.

Finding 1: Schools with well-developed assessment systems align local curriculum, standards, and assessments to state content standards.

Each of the schools visited was aware of increased accountability to educators, policymakers, the community, and the media to meet state student achievement standards. States use large-scale standardized assessments to indicate the extent to which students are achieving the standards, which in turn reflects the impact of a school's educational programming on individuals and

distinct demographic groups. One distinction between schools with well-developed and less-developed assessment systems was their response to rigorous state standards.

Schools with well-developed assessment systems responded in an organized, methodical fashion to rising standards by improving alignment within their local curriculum while ensuring that local standards were consistent with those of the state. This foundation allowed for the identification of local assessments that could indicate student mastery of standards as they were taught. As a result, these schools had balanced assessment systems of which large-scale assessments were one, but not the primary, indicator of success.[2] Aligned local assessment tools were designed to provide feedback to students, teachers, and parents regarding mastery of particular standards relatively rapidly and inexpensively when compared to statewide assessments. Examples of aligned local assessment tools and their impact on the teaching process will be described in greater detail later in this report (see Finding 4).

The process of achieving alignment within these schools was lengthy, arduous, and collaborative. Schools reported spending several years creating time in already-busy schedules, increasing the responsibilities of teachers, and seeking specific professional development to accomplish the task. Typically, this work was viewed as a component of the schools' overall improvement efforts. Schools collaborated with their local school districts or their IEAs throughout the alignment process, and some schools also involved members of the community.

In each well-developed system, alignment began with a detailed analysis of the local curriculum. This analysis required reflecting on several other sources of data, including the state content standards and results from state and local assessments for each subject area and grade, K–12. Therefore, schools with good assessment systems integrated their own history of assessment performance into the alignment and development process. The specific approach to this work varied, but the following four common activities (described in more detail below) tended to drive the process:

1. Curriculum analysis
2. Realignment of the local curriculum
3. Alignment of the local assessment system
4. Reflection upon two data sources: (a) finding from the curriculum analysis and (b) results from state and local assessment in various K-12 subject areas

[2] The large-scale state assessment tool that is referenced throughout is the Michigan Educational Assessment Program (MEAP), begun in 1969. The MEAP has evolved over the years in response to current research in learning, new areas of emphasis in curriculum and assessment, and increased interest in the performance and accountability of Michigan public schools (Michigan State Board of Education, 1999). The latest generation of the MEAP maintains its close relationship to the Michigan Curriculum Framework (MCF) (Michigan Department of Education, 1996), but has shifted its focus away from testing students' minimum, or basic, skills to an emphasis on problem solving and processes—skills Michigan considers essential for students' educational development. In addition to using data from the MEAP, Michigan encourages educators to use and review extensive local assessment data to measure student achievement of core content standards and benchmarks at classroom, school, and district levels (Michigan Department of Education, 1996).

Curriculum analysis. This analysis involved a comparison between the local curriculum and state standards. Schools outlined where state content standards and benchmarks were being taught within subject areas and throughout the local K–12 continuum.

Curriculum maps were used as tools by several schools to help organize the curriculum analysis process. The maps illustrated "gaps," or evidence that certain state standards and benchmarks were not being addressed sufficiently in the local curriculum. They also documented "overlaps," or evidence that certain state standards and benchmarks were being covered in excess from subject area to subject area or from grade to grade.

Curriculum maps also provided data regarding the within-school alignment of curriculum goals as students moved through courses and/or grades, as well as the between-school/district and state/nation alignment of the local curriculum to state and national standards.

Figure 1 is an adaptation from a curriculum map that existed at one of the schools studied. It illustrates in which courses specific arithmetic curriculum strands, content standards, and skills are taught.[3] A series of generic curriculum strands, content standards, and skills are organized in the left-hand portion of the map; the right-hand portion illustrates the sequence of math courses in that particular school. In this example, the definitions of each strand, standard, and skill are not indicated and are not important for our purposes. However, schools that go through the curriculum mapping process would need their curriculum at-hand to make the map meaningful.

Educators can view each academic skill on the map, which is embedded within a content standard and a curriculum strand, and can move across the courses to see in what course that skill is being taught. For example, Skill 1 within Content Standard 2 and Curriculum Strand II is taught in trigonometry and pre-calculus.

Looking at the map in Figure 1, it is clear there are five gaps of specific skills in the math curriculum. Numerous skill overlaps are shown as well, indicating that some topics are not covered at all, while others are covered again and again.

Realignment of the local curriculum. Schools made decisions to adjust the curriculum according to their hypotheses. They advised other educational stakeholders, such as parents and students, of new curricular standards through presentations and/or written documents.

Alignment of the local assessment system. Schools ensured that local assessment tools measured state standards and benchmarks as they were represented in the revised curriculum. In some districts, schools worked together to standardize assessment tools for certain benchmarks at the classroom level. Again, specific examples of local assessment tools will be offered later in this report (see Finding 4).

[3] Curriculum strands are clusters of related content standards representing disciplines or common themes. Content standards are descriptions of knowledge students should acquire in core academic subjects. Skills are descriptions of abilities students should acquire in core academic subjects.

Figure 1. Representation of a Curriculum Map

Curriculum Gap

Curriculum Overlap

Curriculum Strand	Content Standard	Skill	Prep Math	Alg I	Geom	Alg II	Trig	Pre-Calc	Calc	Stats
II	2	1					X	X		
		2				X	X		X	
		3								
		4								
		5				X				
	3	1	X				X			
		2	X							
		3	X							
		4		X						
		5		X			X			
III	1	1		X						X
		2								
		3								X
	2	1								
		2		X		X				X
		3								
		4								X
		5								X
	3	1			X					
		2								X
IV	1	1		X		X				
		2				X		X		
		3				X		X		
	2	1		X		X	X	X		
		2			X	X		X		
		3								
		4							X	
		5		X		X	X		X	
	3	1		X						
		2		X						
		3		X		X				
		4		X		X		X		

Reflection upon two data sources:

a. Findings from the curriculum analysis (i.e., local curriculum gaps and/or overlaps)

b. Results from state and local assessment in various subject areas, K–12

Schools compared characteristics of the curriculum to trends in student performance. This comparison often led to hypotheses regarding the curriculum's effect on assessment results.

In looking at the data in Figure 1, it would be important to understand if gaps in the curriculum were related to low performance on assessment items related to those skills. This finding could lead to the hypothesis that lack of teaching those skills had a negative impact on student performance. However, if students performed well on assessment items related to those skills, it could be hypothesized that no instruction was merited for those skills. Furthermore, there may be a rational, curriculum-related reason that certain topics are covered in several courses; schools should consider assessment results and local goals and needs when reviewing curriculum gaps and overlaps.

In each of the schools with well-developed assessment systems, the process of aligning the curriculum and adjusting the assessment system was done over an extended period of time, and both continued to be reflected upon and modified as local needs were discovered or as state criteria changed.

Schools with less-developed assessment systems were either in the beginning stages of the alignment process or had yet to begin this work. Each of these schools was engaged in ongoing initiatives to improve student achievement and some were involved in initiatives aimed at improving student assessment. Yet because the schools were unable to articulate the connection of their assessment tools to state standards, to their local curriculum, and/or to each other, their assessment efforts did not operate in unison and hence did not compose an assessment "system." For example, one of the schools was in the beginning stages of aligning its curriculum and improving its collection and use of student achievement data. A school-improvement team had been charged with developing a vision for improvement and detailing changes to the assessment system. The present assessment practices were reportedly based on individual teacher interpretations of the standards and curriculum, and a variety of data collection tools (e.g., teacher-made tests, chapter tests from texts, teacher observations). Teachers remarked that they were all "doing their own thing"; that assessment at their school was not systematic; and that decisions about their curriculum and assessment tools were based more on philosophy than on data.

Another school had gone through the alignment process with the aid of many of its teachers and had identified weaknesses in its curriculum (e.g., lack of consistency with state standards and benchmarks or over-/under-representation of topics). They were using large-scale assessment tools to identify high-risk students most in need of interventions, and fine-grained (i.e., individual) tools to intervene and track progress of individuals in accordance with several school initiatives. This school was moving in the right direction, yet seemed stuck at this point. Specifically, they had difficulty conceptualizing how to organize and use their various data sources to help them set goals and improve practices. Moreover, one teacher performing individual interventions with the highest-risk students did not know if teachers were using data generated by her sessions to guide their own work.

Finding 2: Schools with well-developed assessment systems analyze assessment data to monitor student progress.

Schools with well-developed assessment systems used multiple sources of data to evaluate their own longitudinal performance, as well as their performance relative to local and statewide trends. Quantitative analyses of state assessment data went above and beyond reviewing standard reports. Instead, schools found ways to individualize the information and/or format of the reports to meet their specific needs. In some cases this meant disaggregating state assessment data to reflect specific student populations or producing reports on individual student trends in a particular topic strand. Other assessment data, such as locally administered standardized achievement tests, computerized reading and math screenings, or developmental checklists were also included in quantitative analyses.

To make this level of specificity in data analysis possible, the schools used at least one of the following resources:

a. *A "data guru."* This person was a school (principal), district (curriculum director), and/or Intermediate Educational Agency (IEA) leader who had extensive experience in aggregating and disaggregating assessment data and who was effective at building that capacity in others. The leader modeled and facilitated an analytical, data-based approach to educational decision making. He or she encouraged staff to go beyond the role of data consumer and provided the training and time to analyze assessment data and plan interventions accordingly. One principal, herself a data guru, captured the importance of her expertise in this area:

   The role of principal has gone beyond instructional leadership and more toward management. That is part of the reason why many [principals] are struggling. Data use and dissemination of information is a big part of instructional leadership.

b. *A "number cruncher."* If this person was not also the data guru, he or she was typically someone in the district research office who collected, housed, and provided reports on student assessment data, both state and local (e.g., evaluation specialist). In addition, this person responded to school requests to provide individualized reports and provided some professional development for educators. Two major benefits associated with having a research office or evaluator in this position were reportedly:

   a. Evaluators in the research office have the skills and time to respond to specific questions related to student performance. According to one principal, schools need this support:

      I am not a researcher, and I would be lost without these data. They are so beneficial.

   b. By working with individuals who are trained in evaluation and research, educators are able to make better use of assessment data by requesting data that is tailored to their specific needs or reports that are "user-friendly." One principal remarked,

Some of the data I don't care about. They let me personalize the data for my building. They find a way to make [the data] useful...and using [it] rewarding. There is so much good going on with understanding and interpreting data here. States should think in terms of how they can learn from [from this approach].

As a result of continuously analyzing and reflecting upon data in a collaborative fashion, schools had many information feedback loops related to school and individual student performance. Each time they reflected upon data, they had an opportunity to make revisions and improvements. Therefore, none of the four schools with well-developed assessment systems reported relying primarily on large-scale, relatively infrequent assessments to inform their work. Rather, they integrated data from multiple sources to form a cohesive picture of school and student progress.

Schools with less-developed assessment systems were not as advanced in their use of assessment data to monitor student progress. Stakeholders at these schools often pointed to their own lack of technical expertise as a barrier to using assessment data. Schools with less-developed assessment systems either lacked the resources needed to build this capacity in the staff (i.e., no data guru and/or number cruncher) or there was insufficient sharing and communication between those with these skills and those needing the information. Some administrators at these sites admitted to having difficulty poring over and making sense out of standard reports from state assessments, instead relying heavily on individual teacher quizzes, tests, and observations for guidance. Others talked about efforts to have teachers discuss student progress in various subject areas according to standardized rubrics, but had not begun to analyze the quantifiable, local data. Unable to take full advantage of student assessment data, these educators seemed to struggle more with increasing accountability and community attention to student achievement.

Finding 3: Schools with well-developed assessment systems use state assessment data to check the validity of local assessment systems.

Schools with well-developed assessment systems also used state standardized assessment results to validate local assessment tools. To determine the validity of student performance across various tools, outcomes from state assessments were compared with outcomes from locally administered standardized tests and other assessment tools (e.g., computerized reading and math screenings, reading "running records"). This comparison informed educators if student performance at the individual, classroom, and school level was consistent as measured by these different tools. State assessments also provided data that helped schools validate the alignment of their local curricula to state standards as well as the efficacy of their teaching practices.

Schools found ways to aggregate data across several assessment instruments, which helped identify trends across levels of analyses (e.g., student, classroom, school). Following are four of the strategies used by schools with well-developed assessment systems to do this work, including three levels of data aggregation and standardized test item analyses:

1. *Student Profiles.* These profiles allow educators to view an assessment "snapshot" of each child and to identify the degree his or her performances are consistent across various assessment instruments. Student profiles help generate hypotheses about individual children, local assessment tools, the curriculum, and teaching practices (see sample).

Sample Student Profile

Observation: The child performed well in reading comprehension on local assessments and on the district-administered standardized test, but had difficulty with the same skill on the state standardized test.

Hypothesis about the child: The child does well with reading comprehension, but had an "off" day when the state test was administered.

Hypothesis about local assessment tools: The local assessments are not tapping into state standards in reading comprehension.

Hypothesis about the curriculum: We need to devote more time on reading comprehension with this child.

Hypothesis about teaching practices: This child is not learning reading comprehension skills the way we are presently teaching them.

2. *Classroom Profiles.* This method, typically done on a spreadsheet, allows for analyzing performance patterns and generating hypothesis for individual students and groups of students, and across the entire class.

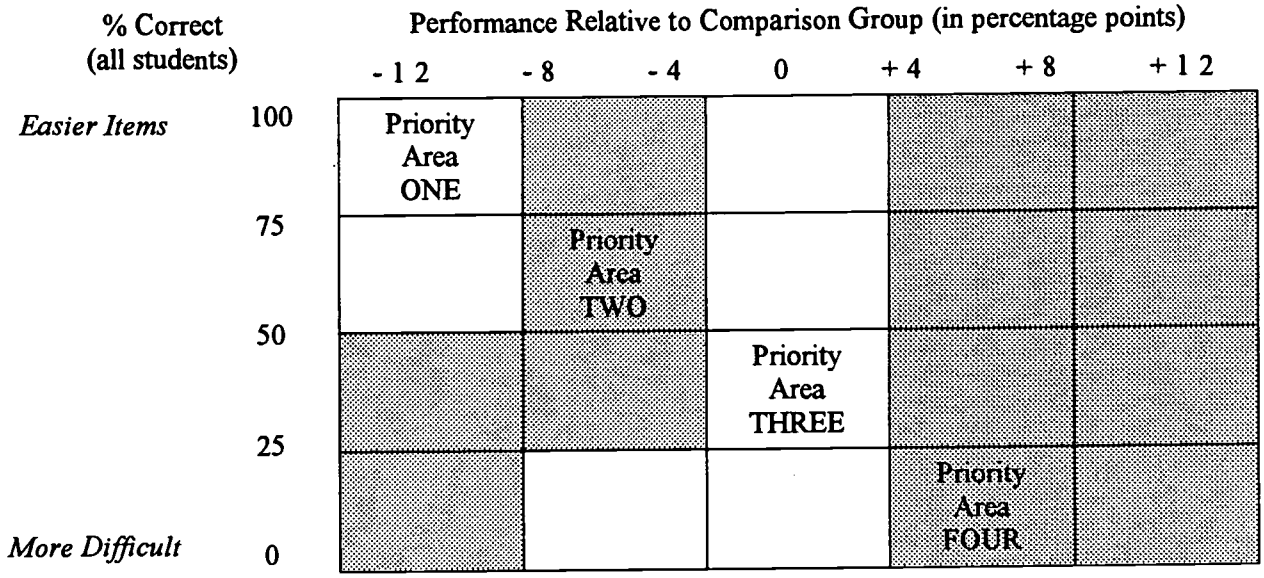
3. *School Profile.* This method, again done on a spreadsheet, allows for analyzing performance patterns and generating hypothesis within grade levels or subject areas. It gives a "big picture" of performance trends throughout the school. For example, one principal reported using these data to identify, then intervene with teachers whose students consistently performed lower on a variety of assessments compared to other teachers. The principal indicated that these data neutralize potentially daunting conversations by providing objective information that can be further probed to identify more specific instructional needs. Further, attempts by the teacher to improve his or her teaching can be evaluated with the same data. The principal acknowledged that many teachers have had negative experiences when using student data because they have been "hit over the head with it." Therefore, the principal used data constructively and attempted to frame these conversations proactively.

4. *Standardized Test Item Analyses.* These analyses involve examining response patterns on specific standardized test items in different topic areas and generating hypotheses about the potential causes of these patterns.

Figure 2 shows an example of an item classification chart, a tool some schools use to analyze their performance on items appearing on their states' assessment. In this example, a school not only analyzed the performance of its students on individual items compared to the performance of students from other schools, it also explored student performance on "easy" (i.e., problems that most students answer correctly) and "hard" items (i.e., problems that most students do not answer correctly).

The item classification chart plots items in four areas according to (a) a school's performance on an item relative to a comparison group (e.g., all schools in the state or district) and (b) the difficulty of the item (expressed as a percentage of all students in the state who answered the item correctly). Items that fall into each of the four areas on the chart have implications for instruction, though those that are located in Priority Areas One and Two tend to generate the most discussion, hypothesizing, and problem solving on the part of participating school staff. For instance, items that are located in Priority Areas One and Two may be further analyzed to examine students' responses to incorrect items. These analyses are typically referred to as "error analyses." For example, when students consistently select or write the same wrong answer, they may have learned the wrong information in class or have similar misconceptions about the nature of the problem, its answer, or both. On the other hand, when students make or select a variety of incorrect responses to a test item, they may have not been sufficiently exposed to the material tested by the item and could be guessing at its answer.

Figure 2. Item Classification Chart



Priority Area One: Items that were easy for the comparison group *but* were very difficult for your school. *Implications:* These items test skills that were probably not taught in your school. Your school should consider focusing on teaching these skills.

Priority Area Two: Items that range from easy to very difficult for the comparison group *but* were very difficult for your school. *Implications:* Items near the top of Area Two involve skills that were probably taught, but not effectively and therefore should be considered for further review. Items near the bottom of Area Two were difficult for the comparison school and your school, indicating perhaps that they are faulty.

Priority Area Three: Items that range from easy to very difficult for the comparison group *but* were relatively easy for your school. *Implications:* Items that appear in Area Three are a lower priority for your school and should be addressed only after those appearing in Areas One and Two have been addressed.

Priority Area Four: Items that range from easy to very difficult for the comparison group *but* were *very* easy for your school. *Implications:* Items in this area suggest that teaching and learning has been effective. If your students tend to score very high on the test, most items will appear in this area. If this is the case, you should prioritize by addressing the most difficult items first.

Educators at schools with less-developed assessment systems discussed having difficulty relating to student performance data from standardized state assessments, which, they felt, did not accurately reflect the complex issues of their student populations. For example, some schools had large and diverse groups of students for which English was not their first language. The emerging ability to read and write English, they felt, introduced a confound for many students that was not sufficiently addressed by large-scale assessments and was not clearly reflected in the results as they were reported. One informant in this situation remarked that large-scale state assessments did not reflect "what we know about our students." That is, the students reportedly performed well on local, school-based assessments but did not perform well on the state, large-scale assessment. The validity of these perceptions, however, was unknown at these schools, as systematic comparisons of student performance across various assessment tools was not being done. Without these analyses, these schools were unable to rely on objective data to demonstrate their assertions or to suggest if specific changes in local assessment tools needed to take place.

Finding 4. Schools with well-developed assessment systems use assessment data to evaluate the efficacy of local curriculum and instructional practices.

Changes to instructional practices occur at the large-scale (i.e., school, classroom) and finer-grain levels. Schools with well-developed assessment systems used universal assessment tools or practices that supported instructional decisions at each of these levels. Like most schools, they used assessment tools at the large-scale level (such as standardized tests) that were familiar to all teachers, yet were administered infrequently. More uniquely, these schools also used or developed assessments at the finer-grained, individual level that were common throughout the school and were administered and analyzed frequently to gauge student progress.

For example, one school created a continuum to document each child's progress through the developmental stages of reading from kindergarten through second grade (See Appendix D). This tool, adopted by the school based on its philosophy of continuous progress, is ideally geared to make fine-grained decisions about individual instructional needs. As children consistently demonstrate reading behaviors, teachers reflect this mastery on the continuum, which follows the child's progress along several dimensions (i.e., comprehension, word identification, attitude) from year to year. Each week, teachers meet to discuss individual student progress along the continuum and to ensure that they are in agreement in its use and interpretation. When the continuum reveals gaps in a student's reading skills, teachers hold conferences with the students and, if necessary, their parents, to problem solve and set instructional goals.

Because reading skills vary and progress along the continuum occurs at different rates among students, teachers at this school also use the continuum for larger-scale decisions about classroom organization during reading instruction. Reading instruction occurs at a consistent time between the lower elementary grades, allowing for flexible groupings of children based on their current reading skills rather than their grade level. Hence, students at the "emergent" level will work with one teacher; those at the "beginning" level will work with another, and so on.

As students raise their skill level, they "flex" (i.e., move) to the more challenging group. Reportedly, this assessment tool has not only allowed teachers to better tailor their reading

instruction, but helped identify instructional needs in teachers (i.e., those whose teaching practices did not support the concept of continuous progress). The success of this tool motivated the third- through fifth-grade teachers to extend the continuum to these grades, improving the continuity and standardization of reading assessment throughout this elementary school.

Teachers in another district worked together to identify two to three standard summative assessments for each district standard in each grade and subject area. After a student has shown mastery of a skill based on the teacher's formative assessment tools (such as observations, classroom tests, and so on), the teacher must use at least one of the districtwide summative assessment tools to "certify" mastery of the standard. Before teachers begin a lesson, they identify which standards are being addressed, what corrective actions they will take with students who do not show mastery of the standards, and what enrichment activities they will provide for students who show mastery more rapidly than others (see the Teaching for Mastery tool in Appendix D). Again, these tools help teachers individualize their instruction according to the different needs of their students and also facilitate flexible groupings within classes based on student's skills. They also ensure some consistency in how children are evaluated within classrooms throughout the district.

Similarly, another school anticipated variable performance among students on formative assessments and designed the following instructional model to manage this:

At appropriate times (e.g., after a lesson, after the completion of an instructional unit), students are assessed on the related standard(s). If the students demonstrate mastery of the standard(s), this is documented with an "M," and the students are moved on to an "enrichment" experience. Students who have not mastered the standard are given an "I" (i.e., incomplete) and are sent through a "reteaching loop," or supplemental instruction. This instruction includes extra work in class, after school, or at Saturday School. Data regarding each student's mastery on various standards are accumulated over the years and form a portion of the student's performance record related to promotional decisions.

Each of these examples suggests a flexible approach to teaching based on individual student needs and an attempt to frequently assess and calibrate instruction to constantly changing student skills. Yet not enough is known about the content of the "corrective" and "enrichment" activities. Often, corrective experiences modify the duration and intensity of instruction. That is, they often reflect traditional educational decisions, such as to spend more time teaching a skill or to practice a skill more. Evidence of innovative instructional practices that work to resolve specific deficits in composite student skills is less common. Further investigation aimed at identifying specific teaching practices within these corrective and enrichment activities at these schools is merited.

Because schools with less-developed assessment systems had not completed the alignment process and did not have staff with the capacity to perform in-depth analyses of their assessment data, they were less able to hypothesize causal relationships between their curriculum, their instruction, their local assessment results, and their state assessment results. As a result, individuals at these schools seldom mentioned the relationship between assessment results and modifications to instructional tools and practices.

Finding 5. Schools with well-developed assessment systems limit the quantity of student assessments to those that are purposeful and aligned.

Assessment costs schools time, effort, and money. Schools with well-developed systems identified which tools were of value to them and discontinued using those that were not. They struck a balance between the depth of the system, the richness of the data, and the cost and utility of the assessment system. The general philosophy driving educators at the school, coupled with findings from extensive analyses of assessment data, such as those previously described, informed their decision-making process regarding the use of individual tools.

At the schools with well-developed assessment systems, educators continuously evaluated the efficacy of their overall system and the contributions of individual tools. When tools were deemed too costly, relative to the utility of the data they offered, they were dropped. For example, several schools reported discontinuing the local administration of certain commercially available standardized assessments. Some were investigating alternative large-scale assessments that were better aligned with local curriculum standards. Others discontinued the practice altogether.

Schools with less-developed assessment systems made the effort to collect data on student achievement, often from multiple sources. Some of these schools were balancing several initiatives sponsored by different agencies, each initiative having its own set of assessment tools. Yet because the schools' approaches to assessment were not systematic and they were unable to effectively analyze assessment data, the payoff for these efforts was unclear. That is, at some of these schools, teachers and administrators reported putting great efforts into assessing student progress but doing little work to analyze, coordinate, and conceptualize the results. Paradoxically, some of these schools may have been "overassessing" their students.

Finding 6. Schools with well-developed assessment systems allocate time for collaboration, reflection, and decision making based on student assessment data.

Teachers frequently report having to juggle multiple responsibilities and to balance competing demands on their time. Nonetheless, teachers at schools with well-developed assessment systems performed much of the hands-on, time-consuming work during the alignment process. They also shared, compared, and discussed assessment data while forming hypotheses and solutions about their own curricular or instructional needs. This typically occurred through a combination of one-on-one meetings or regular small-group sessions throughout the school year. They also shared student assessment data with parents, usually through conferences that they led or by facilitating those that were student led. How, then, did they perform and have time for these formidable tasks?

Teachers suggested that they could not have improved their assessment systems without additional support, due to their lack of experience with these tasks. Much of the support came from local school districts or IEAs and took the form of professional development with teachers. District leaders with expertise in this area provided conceptual frameworks, materials, and guidance for committees of teachers as they performed curriculum gap analyses, defined their local standards, aligned their local curriculum and assessment systems to state standards, and worked to understand student needs. Developing and maintaining these improvements were time-consuming processes.

This work required creative scheduling, allowing time for teachers within schools and throughout districts to meet regularly. The strategies adopted by schools were simple, yet required flexible thinking, a willing staff, and, at times, union negotiation for more accommodating schedules and job descriptions. Table 2 outlines how the various schools managed to create time for teacher reflection and the concrete activities that occurred during that time.

The commitment of districts and/or IEAs to building the capacity of school staff in their use of assessment data was consistent in each of the schools with well-developed student assessment systems. These "capacity-building models" varied slightly from site to site, and the name was not coined as such or visually represented. Yet because of their integral role in the use of assessment data, we have chosen to illustrate an approximation of the models as they were observed in these schools (see Figure 3). You will note that teachers are central in the models. Within the context of receiving training and support from external sources and their instructional leader, teachers do the hands-on work to design their assessment systems and are the primary decision makers regarding curricular and instructional interventions. As a result, participating teachers made great leaps in their own skills, were able to train others, and reported taking pride in their expanded roles. Other positive consequences of this work were reportedly increased communication throughout the system and mutual ownership of the curriculum and assessment system.

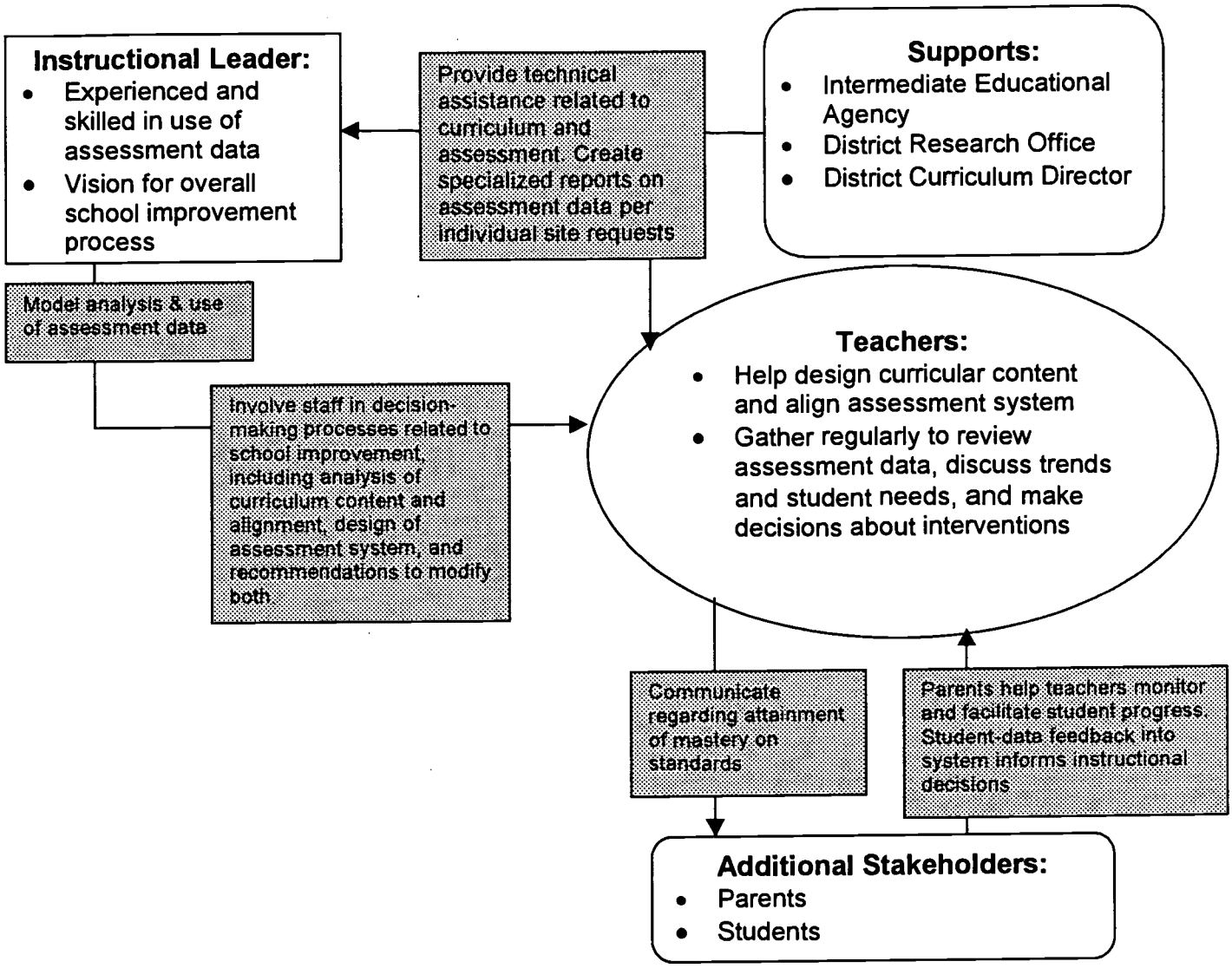
Among schools with less-developed assessment practices, some struggled with the logistics of professional collaboration and lacked the leadership needed to provide professional development targeted toward improving their assessment systems. Several of the schools were unaware of where to turn for this additional support and had not identified ways to make time available to teachers for these activities. Perhaps most important, some of these schools had not yet embarked on the alignment process and therefore had not identified the need for collaboration.

Other schools did embrace teacher collaboration and successfully made time for this activity. However, at some of these schools the activities included only a small group of teachers, not the entire staff. At others, teachers met regularly to discuss student progress and to problem solve regarding student needs. These schools were clearly progressing in the right direction, but lacked the assessment system and related data analyses to inform their decision making.

Table 2. Scheduling Approaches for Teacher Collaboration

	School A	School B	School C	School D
Time and Planning Strategies	1. Once every month, the school day begins two hours later—teachers meet during this time to engage in the activities described below. School makes up this accumulated time by extending the school year.	1. School staff is released early from school once per week for at least 45 minutes. This time is added to other days throughout the week. 2. Entire staff meets once a week for one hour before school. Staff decreased the "nuts and bolts" of the meetings and prioritized work related to assessment.	1. Same-grade teachers meet informally during weekly planning periods and formally every six weeks. To accommodate these planning periods, students in entire grades are sent to "specials" (e.g., gym, art classes). Time is also allotted at regularly scheduled staff meetings. 2. Teachers are released from teaching duties several days each year and are replaced by substitute teachers. 3. Teachers meet with principal up to three times each year.	1. Teachers request time to meet with each other during school hours; substitutes are hired to support this. In addition, teachers meet after school. 2. Teachers meet in "within-grade" and "subject area" teams during their planning hours once per week.
Activities	a. School staff rewrite district standards and realign the assessments they use accordingly. b. School staff continuously reevaluate this work, and discuss and plan changes as needed.	a. Schools use allotted time to align curriculum across grades and with the state standards. This process is driven by student assessment data. b. School staff continuously reevaluate this work and discuss and plan changes as needed.	a. Staff discuss students' progress according to the "developmental continuums" written by school staff. b. Teachers administer individual assessments to students. c. Staff discusses reports on assessment data from district research department.	a. Staff shares knowledge gained from professional development activities that addressed curriculum and assessment. They also discuss student mastery of standards and other outcomes and possible intervention strategies.

Figure 3. Building the Capacity of Data Consumers Throughout the System



Summary and Conclusions

This study describes the general natures and uses of school-based student assessment systems in a small sample of elementary, middle, and high schools. We have relied solely on descriptive analyses and presentation of data obtained during our visits to these schools. In order to make these findings clear to readers of this report, we separated the schools into two groups: those that had well-developed student assessment systems and those whose assessment systems were less-developed. We readily acknowledge that this dichotomy of schools risks oversimplification and some insensitivity to situational hardships at some schools that made it difficult for them to optimally use student assessment data. For instance, some teachers and administrative staff were implementing other school initiatives or reforms and, in so doing, may not have reflected on how student assessments could have assisted them in reaching their goals. Other schools—particularly those in poor and/or multilingual districts—may not have had the staff or resources to simultaneously contend with these demographic challenges while developing and implementing an assessment system.

Nevertheless, the strategy of characterizing the schools in this manner has been useful for bringing some of the broader findings to light. For instance, we found that schools in the well-developed group:

- Were very similar with regard to the general strategies they invoked to align their standards and curriculum to the multiple student assessments they used. These strategies included standardized state and district exams and assessments used in entire schools and in individual classrooms.

- Were similar in the purposes for which they used student assessment results. For example, use of assessment not only informed teachers' and administrators' instructional decision making, but were also used to validate whether all assessment tools were yielding consistent information and remained aligned with state and local goals for learning.

- Were very serious about committing time and resources to developing teachers' and administrators' capacity to reflect upon and use assessment data as the basis for decisions affecting the curriculum and instruction. For example, these schools scheduled time for teachers to discuss the progress of individual and groups of students using evidence obtained from multiple—but coordinated—assessments and, more important, to formulate an instructional response to students' needs.

- Had the attitude that in assessment, "less is more." That is, they maximized the efficiency of their assessment systems, testing as much as was needed, but no more than was necessary.

- Admitted that they would never be "done" designing their curricula and assessment systems. The data gleaned from student assessment served to facilitate "self-correction," or continuous, informed improvement of the system itself.

In contrast, schools in the less-developed group had made less progress in aligning the various elements of an assessment system (i.e., standards, curriculum, assessment tools) and had not articulated a clear purpose for the assessments they used. This lack of progress was compounded by the difficulty these schools had in analyzing and using their current assessment data to conceptualize strengths and needs at the student, classroom, and school level on an ongoing basis. Some of the schools in the less-developed group recognized the benefits of teacher collaboration and were facilitating this process. However, the teachers were not able to inform their work with assessment data to the same degree as teachers in the well-developed schools and did not benefit from the same level of capacity building from experienced, skilled leaders.

Without the availability or understanding of their own assessment data, these schools had difficulty prioritizing which assessment tools to use and which to discontinue. As a result, some were practicing a "more is better" approach to assessment, using too many assessments with no clear connection. In short, these schools lacked what Cizek and Rachor (1994) have called a *planned assessment system*, "one in which each assessment activity that occurs is conducted for a well-articulated purpose; the benefits and beneficiaries of the assessment are clearly defined; the uses of the assessment information are real, tangible, and valued by the users; and the assessments are conducted in an efficient manner, ensuring that assessment time is not wasted gathering redundant information and reducing the time available for instruction" (Cizek, 1995, p. 247).

Our intent for this report was simple: to find out more about how schools use the data they receive from multiple assessments and to identify and describe the factors, conditions, and supports that optimize schools' use of these data. These findings could be used to develop a framework for studying school-based student assessment practices more comprehensively (e.g., in states). This framework may include, for instance, the following five dimensions as a basis for determining (a) the extent to which schools vary along each and (b) how—and how much—each contributes to a coherent, school-based student assessment system:

- Use of shared, coordinated resources
- The degree to which the assessments schools used are aligned with standards, such as those prescribed or recommended by the state
- The degree to which the assessments schools use are aligned with each other
- The extent to which the language of standards for student achievement is used in communication with school staff, parents, and the students themselves
- The extent to which results of assessments are used by teachers and administrative staff as the basis for instructional and other forms of decision making

In the near term, this report can be used by education agencies at different levels of the system to reflect on their own capacities and resources to assist schools in the development of coherent, effective student assessment systems. We conclude this report with some recommendations for how state, regional, and local educational agencies may proceed. These recommendations are organized and presented in four parts: (1) improving networks, (2) involving and supporting teachers, (3) formal professional development for teachers and school administrators, and (4) other resources and assistance to schools.

Recommendations

Improving Networks

- *Develop a state network of Intermediate Education Agencies (IEAs) for exchanging information on the use of state assessment results and other assessments for school improvement planning, curriculum alignment, and developing local assessments.*

  Findings from NCREL's study indicated that some IEAs play an instrumental role in local efforts to improve school-based assessment systems. For example, IEAs provide valuable guidance and technical assistance to schools for mapping local curriculum to state standards and benchmarks, interpreting results obtained from state tests, and developing local student assessments and scoring rubrics and aligning these to state standards and assessments. IEAs might be used increasingly to (a) disseminate information to schools about state curriculum frameworks, standards, and assessments; (b) share local models and strategies for designing aligned, school-based assessment systems; and (c) help identify individuals at the school, district, or IEA level who can serve as data gurus and/or number crunchers for schools, and who can help train school personnel to perform these roles themselves.

  State departments might act as the central node of these networks by providing IEAs with information about the state curriculum frameworks, the state assessment, state content and performance standards, sample scoring rubrics, and general guidelines for interpreting aggregated and disaggregated results. IEAs, in turn, can pass these resources along to schools.

Involving and Supporting Teachers

- *Encourage schools and districts to integrally involve teachers in local efforts to develop and use local assessment systems that satisfy local needs and align with state frameworks.*

  Schools that appeared to be making the most headway in implementing effective school-based student assessment systems were using what may be the most valuable resource available to them: their own teachers. Specifically, two of the characteristics differentiating schools that were identified as having well-developed assessment systems from those that had less-developed ones were: (1) the degree to which teachers were involved in various phases of local assessment work; and (2) the degree to which schools created time for teachers to regularly meet, plan, and discuss the relationship between the assessments they used in the classroom and those used at the school, district, and state levels.

- *Encourage schools to allocate more time—or modify existing schedules—so that teachers may analyze and reflect upon student assessment data, plan revisions to their curriculum and teaching practices, and receive inservice support on how to use student assessment data effectively.*

  Studies of education reform and improvement efforts invariably identify time as a major challenge. This is no exception in the present case. If teachers are going to become more integrally involved in the development of a student assessment system at their school, they

are going to require that either additional time be allocated to this work or that the time that they have in schools is allocated differently to accommodate the need. The schools in the well-developed group appeared to be successful at achieving the latter.

Formal Professional Development

- *Increase the state's presence in the field by providing inservice activities that will assist teachers and school administrators in understanding how the state curriculum frameworks and assessments can help in the development of school-based student assessment systems.*

  It is clear that state curriculum frameworks, content standards, performance benchmarks, and, in some instances, sample test items are typically matters of public domain and, therefore, readily available to many school administrators and teachers on, for example, SEA Web sites or through direct mailings. However, this study suggests that school staff may not be using these resources to develop student assessment systems or in their efforts to generally improve schooling as extensively as intended by the state agencies and commissions that developed these materials.

  Teachers and principals noted during site visits that these documents tended to be too large and complex and were particularly uninviting as *printed* documents. Interviewees said they wanted and needed more examples they could use in the classroom of actual applications of the standards in instruction and/or assessments that align with their state's assessment. Many states do develop supplementary materials or toolkits that can be used by teachers, principals, and district staff to detect discrepancies between local curriculum and state standards, plan subject area instructional units, design classroom assessments, and plan a district assessment system. There may be a need for state departments of education to couple "live" demonstrations of these products in the field whenever possible so that these kinds of materials may be used more effectively in helping schools develop their school-based student assessment systems.

- *Encourage state teachers' colleges to offer courses on the role of assessment in instruction, the analysis and application of assessment data, and the use of state content and performance standards.*

  Principals who participated in this study consistently reported that new teachers, most of whom presumably attended state teachers' colleges, were not knowledgeable about the state standards pertaining to the grade level and/or subject they teach. Nor were they adequately trained in how to assess student achievement and how to apply the results of student assessments to their instruction. SEAs should encourage state teachers' colleges to integrate information about state content frameworks, standards, and assessments into the coursework of preservice teachers. SEAs might also consider developing curriculum units that it could recommend for use in college-level courses for preservice teachers.

Other Resources and Assistance to Schools

- *Provide teachers with samples of state assessment items to assist them in aligning their instruction and/or classroom assessments to state content standards.*

Many of the teachers interviewed wanted to be sure that the instruction they give their students is consistent with state standards in the core subject areas and, therefore, would maximize their students' chances of performing well on future state assessments. Teachers often indicated that examples of items that either appeared on previous state assessments that were no longer in circulation or items that mimicked those used on current assessments would provide teachers with a clear understanding of the kinds of skills students would need to succeed on future state assessments. Furthermore, these items could be incorporated into their daily instructional and assessment practices.

- *Consider returning students' scored state exams, scoring rubrics, and scorers' comments to schools.*

  As an added measure to help teachers validly interpret results and plan effective instruction, SEAs might consider returning to schools photocopies of their students' work, the scoring rubrics that were applied in evaluating the quality of students' responses, and the written comments of scorers. Depending on when state assessments are scored, returned papers can be distributed to the teachers of assessed students during the same school year, or to teachers at the next grade level of the following school year.

- *Explore and develop policies and strategies that address the reporting of English Language Learners' (ELL) results. Also, work with schools with large ELL populations to enhance their interpretation of data from state assessments.*

  Students whose native language is not English are being included with more regularity in standards-based assessments. Higher rates of inclusion in educational assessments, particularly at the state level, can be attributed to policies that are designed to hold schools accountable for the performance of all eligible students. Such policies are rooted in a commitment to ensuring equity in educational opportunity and have been bolstered by recently enacted federal law.[4] It is not clear, however, how data should be reported so that schools, teachers, and families can use the information to improve curriculum and instruction.

  Principals and teachers in schools with high concentrations of students whose first language is not English (English Language Learners) indicated in interviews that they would welcome state performance reports that included aggregate school scores (for each grade level tested) and scores that were disaggregated by students' English proficiency (i.e., "flagged" scores). This practice—coupled with the release of state assessment items, scoring rubrics, and students' scored tests—would enhance educators' understanding of the performance of ELLs in their school and, therefore, assist in planning future instruction for students with special language needs. We also recommend that similar reporting strategies and policies be explored for students with learning disabilities.

[4] On June 4, 1997, President Clinton signed into law the Individuals with Disabilities Education Act Amendments of 1997 (P.L. 105-17). The law introduced new requirements to complement such previous requirements as "free appropriate public education," "individualized educational programs (IEP)," and "least restrictive environment," which have reshaped special education in the past two decades.

References

Baker, E. L., & Linn, R. L. (1997). *Emerging educational standards of performance in the United States* (CSE Technical Report 437). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Cizek, G. J. (1995). The big picture in assessment and who ought to have it. *Phi Delta Kappan, 77*(3), 246-249.

Cizek, G. J., & Rachor, R. E. (1994). The real testing bias: the role of values in educational assessment. *NASSP Bulletin, 78*(560), 83-93.

Council of Chief State School Officers. (1998). *Annual survey of state student assessment programs.* Washington, DC: Author.

Darling-Hammond, L., & Wise, A. E. (1985). Beyond standardization: State standards and school improvement. *The Elementary School Journal, 85*, 315-336.

Glaser, R., & Silver, E. (1994). Assessment, testing, and instruction: Retrospect and prospect. In L. Hammond (Ed.), *Review of research in education, Volume 20.* Washington, DC: American Educational Research Association.

Herman, J. L., Dreyfus, J., & Golan, S. (1990). *The effects of testing on teaching and learning* (CSE Technical Report No. 327). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Herman, J. L., & Golan, S. (1991). *Effects of standardized testing on teachers and learning—Another look* (CSE Technical Report No. 334). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Linn, R. L. (1986). Educational testing and assessment. *American Psychologist, 41*(10), 1153-1160.

Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum, Eighty-seventh yearbook of the National Society for the Study of Education* (pp. 83-121). Chicago, IL: The University of Chicago Press.

Michigan Department of Education. (1996). *Michigan curriculum framework.* Lansing, MI: Author.

Michigan State Board of Education. (1999). *Handbook, Michigan Educational Assessment Program.* Lansing, MI: Author.

National Council on Education Standards and Testing. (1992). *Raising standards for American education.* Washington, DC: U.S. Government Printing Office.

National Education Goals Panel. (1991). *Gauging high performance.* Washington, DC: U.S. Government Printing Office.

Resnick, L., & Resnick, D. (1989). Tests as standards of achievement in schools. In *The uses of standardized tests in American education: Proceedings of the 1989 ETS invitational conference.* Princeton, NJ: Educational Testing Service.

Schafer, D., & Lissitz, R.W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education, 38*(3), 57-63.

Shepard, L. A. (1989). Why we need better assessments. *Educational Leadership, 46*(7), 4-9.

Sokoloff, H. J. (1987). *Dècision-Oriented information systems: Leadership and implementation issues.* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77*(3), 238-245.

Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment.* State University of New York Press: Albany, NY.

U.S. Department of Education. (1994). *Goals 2000: Supporting community efforts to improve schools.* Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. (1991). *America 2000: An education strategy*. Washington, DC: Author.

Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education*, 37-42.

Appendix A: Teacher Protocol

As you know, we are visiting your school today to learn more about the student assessments you use, how you and your colleagues use the results of these tests, and the kinds of consequences that follow from these uses. The Michigan Department of Education (MDE) has identified your school for this study because of the way you use and think about student assessment data as a way of improving teaching and learning. Our organization, NCREL, will use the findings of this study to write a policy paper for the MDE that describes the range of assessment practices that are being used and the school conditions that support effective use of student assessments. Your school will receive a copy of this report.

First, could you tell me a little about yourself? How long have you been a teacher? In this school? What subject do you teach (if applicable)?

the development and/or implementation of school- or district-level assessment systems in the past?

Probes:

- Have you ever had any direct involvement in writing *state* content standards, designing tests, writing items?
- If so, what were your impressions of those efforts? What did you learn from those experiences?

O.K., let's talk a little about the student assessments, or achievement tests that are used in this school...

Types of Assessments

What kinds of student assessments does your school currently use?

Probes:

- Who uses these assessments and their results? How? For what purposes?
- When and how often do you and your colleagues use student assessment data to inform:
  a. Instructional decision making?
  b. Program and/or curriculum evaluation?
  c. School or district policies?

May I take a copy of some of these tools with me?

School Culture Regarding Assessment

Is learning about and using student assessment data a priority in your school and/or district?

Probes:

- In relation to other school- and districtwide initiatives (e.g., technology, parent involvement, school safety), how much emphasis is currently given to using student assessment data more effectively?
- When using student assessment data becomes a top priority in your school/district, what kind of use does it tend to be? Whom does this serve?

Are there any formal or informal mechanisms in place at your school to support the use of student assessments?

Probes:

- For instance, do you meet with other teachers to:
  a. Discuss the student assessments in your school
  b. Analyze results and/or trends in students' scores
  c. Discuss strategies for applying assessment information to instruction, etc.
  d. Collaboratively evaluate student work at the same and/or multiple grade levels?
- How often?
- What tends to be the focus of these discussions? Do you discuss student work in relation to state or other standards and benchmarks?

Utility

Which assessments are most important or useful to *you*? In what ways are these useful?

Probes:

- How do you know when a student has learned something?
- To what extent do you rely on assessments to tell you this?
- Which assessments do you trust most and least? Which guide your thinking the most/least? Why? What kinds of data are you most likely to respond to?

May I take a copy of some of these tools with me?

Do you think that the assessment system at this school works well? Why or why not?

Probes:

- Describe the kinds of school conditions that are necessary to effectively implement and use student assessments.
- What conditions impede successful implementation and use?

Are there any assessments in this school that, if you had your choice, you would rather not use (e.g., those that you are required to use)? Why?

Are there any assessment tools that you wish you had but that you currently don't have access to? Explain.

Alignment

Do the assessments you use link to either state or local standards (or both)?

Probes:

- In what ways?
- Do you feel there is anything missing (i.e., not covered by the assessments)? If so, what?

How do you link what you teach to the different assessments your school uses?

Probes:

- Does this vary depending on the assessment tool?
- Can you give me some examples?

Professional Development

Have there been professional development opportunities to support your use of these assessments or to enhance your general knowledge of student testing? Please describe.

Probes:

- What has been the focus of this training?
- How extensive is it? How often do you and your colleagues participate?
- Has this been useful to you? Why/why not?
- How would you rate the quality of these professional development opportunities?

Do you feel you are sufficiently trained to feel comfortable using the student assessments (and their data) implemented at your school?

Probe for details, examples.

Cost / Benefit

We've talked about your perceptions of the utility of assessments and the professional development and time required to use them effectively. How do you assess the bottom line? Are the benefits of these assessments worth the time and effort you put into using them?

Closing Question

Before we conclude, I want to give you the opportunity to tell MDE what you think is the most important thing they should know about how your school uses student assessment data, the challenges it and *you* as a professional face, and how the Department could support your efforts?

Appendix B: Administrator Protocol

As you know, we are visiting your school today to learn more about the student assessments you use, how you and your colleagues use the results of these tests, and the kinds of consequences that follow from these uses. The MDE has identified your school for this study because of the way you use and think about student assessment data as a way of improving teaching and learning. Our organization, NCREL, will use the findings of this study to write a policy paper for the MDE that describes the range of assessment practices that are being used and the school conditions that support effective use of student assessments. You will receive a copy of this report.

First, could you tell me a little about yourself? How long have you been a principal/coordinator? In this school?

Have you ever had any direct involvement in the development and/or implementation of school- or district-level assessment systems in the past?

Probes:

- Writing **state** content standards, designing tests, writing items?
- If so, what were your impressions of those efforts? What did you learn from those experiences?

O.K., let's talk a little about the student assessments or achievement tests that are used in this school...

Types of Assessments

What kinds of student assessments does your school currently use?

Probes:

- Who uses these assessments and their results? How? For what purposes? How often?
  a. Instructional decision making?
  b. Program and/or curriculum evaluation?
  c. School or district policies?
- May I take a copy of some of these tools with me?
- Is there a written school or district plan for using student assessment results? If so, may I take a copy with me?

Could you briefly describe the process that led up to the choices your school (or district) has made in terms of the assessments it uses and how it uses assessment results?

Probes:

- What kinds of assessments have you used in the past?
- How have the assessments you've used—and the purposes for using them—changed over the years? Why did they change? What have you learned?
- Are you planning any changes in the near future with regard to student assessment? If so, what are they and why are you making them?

Alignment

Has your school or district made any efforts to align the multiple student assessments it uses to each other and/or to content standards specified by either the district or state? If so, can you briefly describe this process and the status of the work?

Probe:

- How do you know when assessments are aligned either with each other or with content frameworks/standards?

Professional Development

Have there been professional development opportunities to support your staff's use of these assessments and/or to enhance their general knowledge of student testing? Please describe.

Probes:

- What has been the focus of these efforts?
- Have you had opportunities for professional development in using student assessment information in ways that are distinct from how your staff might use assessment information?
- How would you rate the quality of these professional development opportunities?
- Have these opportunities been useful to you and your staff? Why/why not?
- Do you feel you are sufficiently trained/competent to use the student assessments (and their data) implemented at your school?

Costs/Benefits

How costly are the student assessments your school uses in terms of:

- Actual expenditures for materials and professional development?
- Aligning assessments with each other and/or district/state content standards?
- Time away from instruction?
- Other?

Do you think the benefits of these assessments are worth their costs? Explain.

School Assessment Culture and Support

Is learning about and using student assessment data a priority in your school and/or district?

Probes:

- In relation to other school- and districtwide initiatives (e.g., parent involvement, school safety), how much emphasis is currently given to using student assessment data more effectively?
- When using student assessment data becomes a top priority in your school/district, what kind of use does it tend to be? Whom does this serve?

Are there any formal or informal mechanisms in place at your school to support the use of student assessments?

Probes:

- For instance, do teachers meet with you and/or other teachers to:
  a. Discuss the student assessments in your school?
  b. Analyze results and/or trends in students' scores?
  c. Discuss strategies for applying assessment information to instruction, etc.?
- How often?
- What tends to be the focus of these discussions?

Utility

Which assessments are most important or useful to *you*? In what ways are these useful?

Probes:

- How do you know when a student has learned something?
- To what extent do you rely on assessments to tell you this?
- Which assessments do you trust most and least? Which guide your thinking the most/least? Why? What kinds of data are you most likely to respond to?

Do you think that the assessment system at this school works well? Why or why not?

Probes:

- Describe the kinds of school conditions that are necessary to effectively implement and use student assessments.
- What conditions impede successful implementation and use?

Are there any assessments in this school that, if you had your choice, you would rather not use (e.g., those that you are required to use)? Why?

Are there any assessment tools that you wish you had but that you currently don't have access to? Explain.

Closing Question

Before we conclude, I want to give you the opportunity to tell MDE what you think is the most important thing they should know about how your school uses student assessment data, the challenges it and *you* as a principal face, and how the Department could support your efforts.

Appendix C: School Summaries

Elementary School A

Variables	
Assessments	• Emphasis on district K-2 continuum ("continuous progress") for reading, writing, and math development. School faculty chose to extend K-2 continuums to 3rd to 5th grades. • Classroom assessments that were used to provide data for continuum include: Reading Recovery, "Literature Circles," in-class writing portfolios, etc. • MEAP data analyzed in great detail but test is not primary catalyst for curricular and instructional decisions.
Uses	• Plan instruction • Identify/request/analyze additional data needed • School Improvement planning • Smooth transition from grade to grade
Status of Alignment	• Continuums based on theory loosely coupled with MI Standards (include some state standards, but not all). Discussion regarding including more state benchmarks in future versions.
Processes	• Schedule allows for same-grade teachers to share planning times 1 time a week informally, 1 time in 6 weeks formally regarding student progress on continuums • Staff meetings prioritize discussion about assessment data • Principal meets with each teacher 3 times a year to discuss reports from District Research Department (including Standard MEAP reports, school-requested analyses regarding MEAP, and continuum data for classrooms and individual profiles) • All teachers get several days a year (K, 8 days; other elementary grades, 4 days) when substitutes hired for classrooms and teachers devote time to perform individual assessments with students. • Teachers collaborated on development and use of local assessments (e.g., continuums).
Professional Development	• Initially, district guided workshops (group level) regarding development and use of continuums. Currently, individual assistance provided. • Professional Development at group level not emphasized, but individual mentoring regarding use of data provided, particularly by school principal.
Additional Comments	• District Research Department takes active role in collaborating with schools regarding disaggregating data and producing custom reports. • MEAP part of overall district/school assessment plans, but local assessments (i.e., continuums) given most credibility. • Teacher involvement in developing continuums and in decision making about student data engenders sense of ownership and empowerment.
Advice to State	• Focus on process, not products: Assist schools to find adequate planning time to use assessment data and deemphasize "products" that relate to distilling data. • Integrate concepts of assessment/use of data into teacher mentoring and preservice courses for educators. • Report data in simplified, user-friendly form. Allow schools to personalize reporting.

Elementary School B

Variables	
Assessments	• Formative assessments including Developmental Reading Assessment (DRA), running records, teacher tests, labs, projects, etc. Teacher chosen. • Summative assessment: Critical learnings assessments standardized throughout district K-12 for all subject areas and organized in binders; 2 or 3 tools selected by committee for each critical learning, rubrics for scoring; minimum of one must be used to validate that student has met critical learning (district standard). • MEAP viewed as a "report card" and an additional source of data. More emphasis on local (district) assessments.
Uses	• Results on formative assessments inform teacher regarding student readiness for summative assessment. • Results from summative assessments inform teacher regarding student's mastery of critical learning; used to report to parents (with developmental rubric) in lieu of traditional letter grades. • MEAP data will be used to validate alignment of district standards with state standards.
Status of Alignment	• Critical learnings reflect district standards; summative assessments intended to tap these standards. • District standards are aligned with MI standards but reflect specific local needs.
Processes	• QSP: Quality Schools Process embraces site-based decision making. • Teachers, community, and board took significant responsibility in designing district standards and creating summative assessments for critical learnings. • Making time for this: scheduled once a month for 2 hours, delayed-start school day to redesign standards/align assessments. Made up for time with extension to school year • Parents/students given critical learnings for that subject/grade at beginning of year. Reported on 3 times a year to parents. • Teachers took MEAP to gain insight: How do we need to teach so that students are able to answer these questions?
Professional Development	• Related to standards, critical learnings, creating/using assessments, rubrics, etc. • Combination of in-house ("train our own") and external training (send individuals, they train others)
Additional Comments	• Strong leader has lead total "overhaul" of system within past year • Philosophies: continuous progress, mastery learning, site-based decision making • Focus: locally defined outcomes aligned with state standards; locally created assessments aligned with critical learnings
Advice to State	• Offer incentives (earmarked PD funding, grants) to districts to offer/take advantage of professional development (PD) around standards, assessment. • Demand feedback regarding how PD funding is used. Hold districts accountable. • Be more visible throughout the entire state in PD efforts, particularly for administrators

Elementary School C

Variables	
Assessments	• Investment in individual classroom assessments, informal and formal: e.g., journaling about students during observations, Reading Recovery, Running Records, pre/post testing before/after units taught. • Districtwide: Terra Nova test (3rd, 6th, 9th). District feels it is relatively aligned with their curriculum. • Schoolwide: STAR Standardized Test for Reading Achievement: computerized administration 4 times/year • MEAP: Not an "end all" source of data, but a check-point for validity: Are our students learning well? Also valued as a common measure throughout the state to allow for school-to-school and school-to-state comparisons.
Uses	• Classroom assessment information integrated with other standardized assessment data to help teachers adjust instructional strategies for different individuals/skill levels within classroom. • MEAP: use item analyses, content strand analyses to help teachers digest trends, suggest adjustments in teaching/curriculum. • Principal also uses MEAP data as a reflection of teacher efficacy. Will counsel teachers who are consistently under performing others in same grade level. Data makes this conversation objective.
Status of Alignment	• Terra Nova perceived as more aligned with district curriculum (than CAT). • Districtwide effort to standardize classroom assessments within grade level and ensure alignment with standards. • Major changes started 2 years ago to realign curriculum districtwide and define "core curriculum components."
Processes	• Teachers took MEAP to familiarize themselves with the test and its demands. • ISD provides MEAP item analyses to district; district trains administrators to use. • Principal provides a class list and all data available on each student (e.g., MEAP, STAR, Terra Nova, etc.) to teachers. He or she looks with teacher for patterns of difficulty in students across assessment types. Use to guide interventions. • Many teachers work with interns from nearby university all year, which frees some time for more individual assessment. • At district level, committees composed of teachers and district leadership compare district curriculum with state benchmarks, then relate assessment back to the curriculum. Will roll out to other teachers in the district. • Making time for this: (1) prioritize use of assessment data at weekly staff meetings, (2) early release time 1 time a week for 45 minutes for teacher collaboration/reflection. Time is added to other days.
Professional Development	• District puts on workshops regarding tying assessment to curriculum and to teaching. • Teachers work with item analysis, data analysis (with principal). • Each teacher has personal PD plan and collects data regarding development. This is reviewed periodically with principal..
Additional Comments	• Good collaboration among school, district, ISD regarding use of standards/aligned assessment data to guide student intervention. • Strong school leadership: high standards; holds teachers accountable; has extensive background in data use and staff development; models and counsels teachers to use data to improve; proactive, not reactive. • Philosophy regarding assessment data: Use it to inform and guide, not to bully teachers or accuse them of not doing their job. • Expectation: teachers gather data, principal analyzes and distills data and then communicates with teachers about results.
Advice to State	• More PD regarding principles (i.e., reliability, validity) of assessment for teachers • Need menu of assessment tools for teachers to draw from, which would create standardization throughout state. • Condense/graph MEAP results for easier communication • More PD regarding how to USE data, how to create time to reflect on data. Make appropriate decisions based on data.

Elementary School D

Variables	
Assessments	• Daily Progress Reports for students identified as "falling behind" in work: written by teacher, checked and signed by principal, checked and signed by parent, returned to teacher. Continues until improvement is noted: no specific criteria to begin/discontinue this intervention. • Student portfolios with teacher reading and writing assessments (Title I teacher administers different district-identified assessments for different grade levels) plus student-chosen representation of best work. • Individualized reading assessments, Running Records, ESL literacy: Demand much time, but perceived as most valuable assessment tools. • Woodcock-Munoz for LEP: large LEP population at this site. • Terra Nova: new choice by the district.
Uses	• All students are at-risk at this site. Use Terra Nova to "triage" population: identify MOST at-risk (i.e., <=40th percentile) and provide interventions to them first. • Terra Nova used to supplement MEAP info • MEAP: How are different populations (ethnicity, gender, etc.) doing on particular test items? Because large ESL population, use item analyses to identify common vocabulary use in frequently missed questions: what words are throwing these students off? Use info to tutor students; use different performance levels on MEAP (High, Medium, Low) to organize future testing groups (i.e., test High group separately from Low because High group tends to finish more rapidly and may discourage Lows)
Status of Alignment	• MI standards and district texts: extent of alignment has be identified • Alignment of instruction to standards: K-7 teachers have own curriculum binder and codes for standards: must refer explicitly to standards as they addressed in lesson plans. Principal enforces. • Insight regarding alignment of individualized tests; district-chosen test (Terra Nova), MEAP not clear, confusing
Processes	• SIP met last summer to align standards to texts: Realized assumptions that these were aligned were often wrong. • Curriculum mapped last summer K-3 (4-5 will be next). Identified gaps and overlaps, realigned and organized with standards: now timing instruction of topics to align with when topic addressed on the MEAP. (Communication of this new strategy is under way.) • Started last year: District-level committee gathers 4 times a year to evaluate student writing samples against MI state rubrics: used data to publish manual with examples and this district's writing/corresponding ratings from rubric. This year comparing samples to last year: evaluating progress of students and district overall. Use to discover what types of instruction push student writing skills.
PD	• Traditional inservicing: some hands-on beginning work with alignment.
Additional Comments	• Conceptualization skills need development and guidance: How do we organize all of these data, see how they all hang together, use them to set goals and priorities, and then apply them to practice? • Competing initiatives, how do they synthesize? Many improvement programs are going on at once: Need help fitting assessment data from these together or creating priorities among the initiatives. • Title I Teachers doing many of the individual assessments: not sure if teachers are using this information for own instruction.
Advice to State	• Tailor PD efforts regarding use of assessment data to schools at different stages of the learning curve. This school more "novice" with more basic needs: Per principal, "We are an emerging district"; others are further along and will have different needs, challenges, and questions. • Help schools with high LEP population understand how to tie standards to these students at various proficiency levels and how to better apply results from MEAP to the instruction of these students.

Elementary School E

Variables	
Assessments	• Teacher-made tests including verbal/oral assessments • Teacher observation • End-of-chapter tests and assessments provided by curriculum packages (e.g., phonics) • Stanford Achievement Test (v.9) in grades not tested by MEAP • MEAP
Uses	• Teacher-made tests, observations, and end-of-chapter tests are used to monitor individual student's progress, guide instruction. • MEAP results are disaggregated and used to look for "holes" in students' learning; process is not systematic. • SAT used to get annual data on students; assists K-3 teachers in getting all students to read by 3rd grade (school goal).
Status of Alignment	• School has adopted (a) curriculum developed by ISD, (b) state curriculum frameworks with some flexibility for local interpretation, and (c) elements of known/exemplary programs in other school districts. • State framework will be used to drive curriculum mapping. • School staff plans to conduct the mapping themselves; "engenders ownership" (4th-grade teacher). • Progress on mapping has been slow; superintendent estimates that this may take "a couple of years."
Processes	• Teacher grade-level teams meet informally. • School Improvement Team—composed of 5 teachers, parent, principal, and school board member—responsible for developing and implementing "school vision." • School is in beginning stages of developing a "school portfolio" that will reflect—and guide the implementation of—the school's vision for improvement. School portfolio will include a framework for improving school practices in the area of collecting and using student achievement and data systematically and effectively.
Professional Development	• Professional development around assessment not emphasized. • School receives much of its professional development support and resources from ISD.
Additional Comments	• Sentiment for self-determination is strong. • School does not over-assess; limits the number of assessments to those they believe to be most instructionally useful. MEAP is important, but it is not at the center of the school's plans for improving. • School in transition: new leadership, developing school vision and ways to involve more staff in decision making. • A coherent, systematic plan for student assessment not yet in place. Teachers say that the data are not driving the process and they are "all kind of doing our own thing…it's not systematic."
Advice to State	• Return students' actual tests so that teacher can, in combination with the results, used them to plan instruction. • Develop flexible, inquiry-based state database that schools/teachers can use to (a) disaggregate MEAP data and generate reports that address their specific needs, and (b) assist in the alignment and/or modification of curriculum.

Middle School A

Variables	
Assessments	• Classroom assessments that were developed to align with ISD-developed curriculum • Student portfolios that include writing samples; writing across the curriculum is emphasized • Individual student "cards" that contain "snapshot" information about students • No district standardized test—ITBS dropped because school wasn't using these data • MEAP
Uses	• Classroom assessments and portfolios emphasized in this school for planning instruction, grouping students • Student cards are written and used in May of each year to assist teachers at next grade level to develop class lists (rosters) based on a mix of student strengths/weaknesses. • MEAP results used for cohort and item analyses—items are plotted on matrix that is used to compare school performance to county's.
Status of Alignment	• ISD-developed curriculum drives most of what is done (MEGOSE in science). • Considerable work done to rewrite and align this curriculum to MEAP benchmarks (e.g., "gap analyses"); work performed by teachers, led by staff leaders who were trained by ISD and/or had extensive in-servicing around gap analyses.
Processes	• Considerable amount of time allotted for staff planning. • Parents and students are aware of standards—students' performance is reported to parents according to standards and benchmarks • Decisions about whether to "enrich" or "reteach" students on content are made after each lesson, unit, and standard is assessed—this is standard practice in the school. Students who master standard move on to enrichment activities. Students who don't are "looped" back for more instruction—students go to Saturday or summer school if standard is not mastered. • Teachers are represented in committees that are responsible for decision making at various levels of the system (i.e., school, district, school board); this enhances consistency in language/goals of school and district improvement plans, North Central Accreditation process, and process of aligning MEAP with ISD-developed curriculum. • School board very active in assisting school to identify goals for improvement.
Professional Development	• Major commitment—time and money—made to professional development; 7 days per year per teacher. These resources are directed away from "bells and whistles" programs and staff-development. • Staff regularly attend inservice training sponsored by their ISD and are occasionally sent across the country to be trained by experts. • Teachers are expected to train each other based on their inservice experiences and/or committee work (e.g., curriculum gap analyses).
Additional Comments	• Have pared down the number of student assessments they use to only those that the staff feel are essential. • Staff are trusted in making decisions; have a sense of ownership in school improvement initiatives and are professionally gratified. • Role of ISD larger than district office; ISD considered innovative. • Staff considered by leadership as most important resource.
Advice to State	• Need to be able to break down student MEAP performance data according to schools' inquiry needs. Would increase the power of the data. • Electronic network to house all data in one place—have at fingertips of school teachers, administrators. • New teachers need to be more adequately trained on state standards, the MEAP, and use of student data to drive decisions.

Middle School B

Variables	
Assessments	• Student Learning Styles Inventory (i.e., visual, auditory, kinesthetic) • Classroom pre- and post-unit tests • Pretests at beginning of year and pre-MEAP writing prompts • "Profiles in Writing"—identifies quality writing and analyzes attributes within grade levels, across curriculum • District assessments include (a) California Achievement Test (CAT) at 3rd and 6th grades ("off-MEAP" grades), and (b) teacher observations/recommendations (e.g., work ethic, creativity, etc.) • MEAP
Uses	• Learning styles, pre-/posttests, and classroom assessments used to modify instruction. • CAT assessments in reading, math, science, and social studies used to rank 3rd graders and to identify gifted and special-needs students for pull-out programs. • MEAP, classroom assessments, and teacher recommendations used to change groupings of students (i.e., students may "flex in or out" of programs based on these measures.
Status of Alignment	• Class syllabi are developed according to state standards and benchmarks.
Processes	• Teacher teaming within grade level: They meet once a day for 40 minutes; interdisciplinary. • Chair of teacher team is a liaison between district and her colleagues. • Within subject area, teachers meet once a month; use this time to set goals; discuss/plan curriculum, professional development, evaluate strategies (e.g., writing across curriculum); trade notes/data on individual students; ensure that they are adhering to and implementing North Central Accreditation (NCA) guidelines. • Teachers participate in Writer's Interactive Network—uses "Profiles in Writing" to support student writing. • School uses MicroSage, a software program that manages districtwide database of student performance on MEAP and district assessments, attendance, medical data etc.; links to budget, expenditures; and generates individual or group reports.
Professional Development	• Not limited to assessment topics, but many teachers choose to attend conferences/workshops related to MEAP, state standards, and NCA • One day for MicroSage training
Additional Comments	• Focus at this school is on adhering to and reaching for state standards as opposed to focusing on performance on MEAP itself. • District is goal-oriented: has goals and ideas about how to get there; engages in self-reflection. • Teachers report improved relations among colleagues in terms of sharing information about students, resources, and feedback. Benefit is in seeing linkages between grades and schools as well as better understanding of standards, strategies, use of data.
Advice to State	• Direct communication to teachers about standards, MEAP scoring, and benchmarks; information directed to school administrators doesn't always get filtered down to staff. • Distribute more examples of scoring rubrics to schools. • Teachers need "global view" of how students are—according to state standards—to progress from year to year. How does an 8th-grade teacher influence student performance in the 10th grade? • Provide for voluntary teacher seminars that are relevant to needs; give teachers opportunity to share ideas.

Middle School C

Variables	
Assessments	• Starting to use portfolios: Within-grade teacher team adds work, student chooses work, and then student leads a portfolio conference with parents. • School is investigating alternatives to MEAP (e.g., CAT). Feels CAT is "parent-friendly" in its reporting (i.e., more verbal, less quantitative). • Incongruence between district and school regarding MEAP. District sees it as the centerpiece for assessment and understanding achievement. School (administration, staff) does not put a lot of credence in MEAP results. • School rationale regarding MEAP • The test (not the standards) is not realistic in its expectations for what students are able to do. • "MEAP does not reflect (i.e., is not consistent with) what we know about our kids." • Questions are confounding because they involve difficult vocabulary (i.e., cannot tease apart their knowledge of that standard from their ability to understand the wording of the question).
Uses	• School level: MEAP used to help identify teaching goals and needs, but emphasis more on NCA standards. • District: MEAP used to look at longitudinal trends, and whether money is allocated correctly to particular subgroups.
Status of Alignment	• MI Standards not a focus. Instead, school is focused on NCA goals, goals for specific grants.
Processes	• Within-grade teams meet daily for 50 minutes; teams of "Encore" teachers participate in portfolio meetings to analyze and plan according to various data sources, compare notes on students, identify student strengths and weaknesses, set goals, and design action plan. • Item analysis of MEAP on within-grade teacher committees. • District has been asked by State to develop objectives regarding how MEAP data will be used in all schools.
Professional Development	• ISD offers PD for teachers related to MEAP (how curriculum should relate to standards, which are tested on MEAP). • DataMetrics software to help disaggregate MEAP data: but do school personnel know how to use it? • Curriculum director works with principals to help comprehend MEAP data.
Additional Comments	• School has site-based management: helps explain some inconsistencies between school and district. • Lack of credibility of MEAP may be related to school's novice status regarding knowing how to efficiently use the resulting data and also to emphasis on goals that may not be consistent with state standards (therefore, what is being measured in-class may not be aligned with what is being measured on the MEAP). • Focus on NCA and other funders because qualitative data and technical assistance are being provided by these sources. These sources provide the guidance that is perceived as not being offered by the state.
Advice to State	• Form a coalition of teachers from a variety of school populations to help revise some questions on the MEAP (i.e., those that are frequently missed, or consistently missed by certain populations). • Inservice schools regarding MI standards, their purpose, how to use them effectively to drive curriculum, instruction, assessment. • Recommend alternative assessments (including those that are computer-based) for schools that are aligned with MI standards. Encourage software developer and publishers to develop these aligned tools. • Encourage universities to train preservice teachers how to align instruction to standards, how to interpret assessment data, how to use interpretations to inform instruction. • Hold teachers accountable for teaching according to the standards. This must be universally enforced. • Provide simple screening assessments (e.g., in reading and math) in various forms (to avoid practice effects) that are aligned with MI state standards. Teachers can use assessments at the beginning of the year with new students and recheck with a different form throughout the year. Test correlation of this tool to MEAP.

High School A

Variables	

Assessments	• Individual writing interviews (Jr.)
	• Writing folders
	• ELA classroom assessments—tied to NCA goals
	• CTBS (Grades 2, 3, 4, and 8); Terra Nova (Grade 6), Harcourt-Brace ELA assessments (K-8); Macmillan pre-post tests; all/most tests administered one grade level prior to MEAP
	• MEAP:
	  • Frustration regarding how writing portion is scored (i.e., quality of scorers); also, hard to give students feedback on writing when school is not offered specific information regarding rationale for scores.
	  • Standards perceived as being too difficult. Students at this school, many of them high achievers, are not performing as they would be expected on the MEAP.

Uses	• Emphasis on writing assessment
Status of Alignment	• School NCA goals are aligned with state standards. • MEGOSE provides centerpiece for science curriculum. • Conscious effort to design school-based assessments that are consistent with state standards.
Processes	• Each school has a coordinator responsible for updating staff regarding changes to MEAP: "We want to make sure we teach what is tested." • One teacher is freed up for a semester to align curriculum across high school grades, according to MI standards. • Beginning to consider aligning middle school curriculum with high school curriculum across subject areas. • Movement toward creating a "data culture": with a variety of assessment data at the fingertips of administrators. • Teachers serve as members of curriculum and assessment committees.
Professional Development	• Not heavily emphasized, but available to teachers
Additional Comments	• Classroom and MEAP assessments are two halves of a whole. • Concerns that changes in the MEAP are causing a downward thrust of topic areas that are developmentally inappropriate (i.e., too complex) for younger students. • This district has financial resources to free-up/hire personnel to analyze data, work on mapping curriculum, etc. SEA may need to compensate with additional training to schools without these resources.
Advice to State	• Advise districts how to align curriculum to state standards. • Provide richer information on MEAP reports (e.g., item analysis, scores by strand) to better inform instruction.

Appendix D: Teaching for Mastery—Teacher Outline

<table>
<tr><th colspan="3">TEACHING FOR MASTERY</th></tr>
<tr><td colspan="3">The learner will:</td></tr>
<tr><td>Standards to teach/assess</td><td colspan="2">3.4 Demonstrate knowledge of nonfiction by:
a. Identifying author's purpose.
b. Determining topic, main idea, and supporting detail.
c. Comparing and contrasting.
d. Determining what is fact or opinion.
e. Locating information in appropriate reference materials (dictionary, encyclopedia, atlas, almanac, newspaper)</td></tr>
<tr><td colspan="3">Vocabulary</td></tr>
<tr><td>Prerequisite</td><td colspan="2">New</td></tr>
<tr><td>(vocabulary needed prior to instruction)</td><td colspan="2">(vocabulary to be introduced during instruction)</td></tr>
<tr><td colspan="3">Instructional Plan</td></tr>
<tr><td>Information/Concepts</td><td>Procedures to Follow</td><td>Instructional Sequence</td></tr>
<tr><td>(main instructional concepts)</td><td>(preteaching, instructional steps, materials needed)</td><td>(organize sequence in which information is introduced)</td></tr>
<tr><td colspan="3">Assessment</td></tr>
<tr><td>Correctives</td><td>Enrichment</td><td>Summative</td></tr>
<tr><td>(corrective action for students that do not master concepts)</td><td>(enrichment activities for students that show mastery of concepts before or after the lesson)</td><td>(identification of which districtwide assessment tool(s) will be used to certify mastery of concepts)</td></tr>
</table>

TM032506

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").