# Important Features of State Assessment Systems from the Local Perspective: Interim Report

CSE Technical Report 472

Brian M. Stecher and Sheila Barron
CRESST/RAND

Hilda Borko and Shelby Wolf
CRESST/University of Colorado at Boulder

July 1997

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-6511
(310) 206-1532

The work reported here was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

# IMPORTANT FEATURES OF STATE ASSESSMENT SYSTEMS FROM THE LOCAL PERSPECTIVE: INTERIM REPORT

**Brian M. Stecher and Sheila Barron**
**CRESST/RAND**

**Hilda Borko and Shelby Wolf**
**CRESST/University of Colorado at Boulder**

**Abstract**

This is the first report from a five-year study of the classroom level effects of state assessment reforms. The first year examines changes in the teaching of writing and mathematics in Kentucky elementary and middle schools as a result of the Kentucky Instructional Results Information System (KIRIS) and the Kentucky Educational Reform Act (KERA). This report presents a framework for describing the effects of assessment reforms at the classroom level. The authors describe eleven features of state assessment reform that are salient at the classroom level and compare the perspectives of classroom teachers and researchers on these features. The list of features should alert researchers to important aspects of assessment reform, and the comparison of the two perspectives should help both groups think more effectively about these issues.

## Introduction

This is the first report from a multi-year study of the effects of state assessment reforms on classroom practices. The research focuses on changes in curriculum, instruction, and assessment at the classroom level and the structures that support or hinder local change. In 1996-97, the study examines the impact of the Kentucky Instructional Results Information System (KIRIS), which was created by the Kentucky Educational Reform Act of 1990. In subsequent years, a second state will be added to the study, permitting comparisons of the local effects of two different state assessment reforms.

We recognize that state assessments can serve many purposes, but we are focusing on those designed as levers for instructional change. We use the shorthand

label "assessment reform" to refer to those assessment system changes that are linked to broader educational reforms.

Similarly, we recognize that there are other questions about assessment reforms that are of interest to policymakers. Our focus on the impact of state assessments at the classroom level does not negate the value of research on the development and features of state assessment systems (Aschbacher, 1991; Bond, et al., 1995), the implementation of assessment reforms (Koretz, et al., 1994), or the technical quality of assessment results (Hambleton, et al., 1995; Klein, et al., 1995; Koretz, et al., 1996). This study focuses specifically on aspects of state assessment that are relevant to the classroom or "local" level.

In this paper we present a framework for studying assessment reform that compares the viewpoints of teachers and researchers toward the impact of assessment reform at the classroom level.[1] Our experience suggests that teachers and researchers have different interests, so they look at the question of local effects through different lenses. Teachers are concerned that students in their class and their school receive the best education possible. Therefore, they may focus on the effects of assessment reform on their day-to-day classroom responsibilities. In contrast, researchers are interested in general statements about the efficacy of state assessment reforms and how such reforms might be improved. Consequently, researchers may think more about the commonality of teachers' concerns and experiences, the pattern of effects across classes, and the trends in impact over time.

The members of the research team believe that the differences between the teachers' and researchers' perspectives noted above are more than merely academic, they can make a difference in the effectiveness of both groups. Teachers who do not think about assessment reform beyond the walls of their classroom or school may adopt solutions that are not as effective in the long term as solutions derived from wider information. Researchers who overlook questions of immediate import to individual teachers may draw invalid conclusions about assessment reforms that ignore issues of concern to teachers. We believe the delineation of features will alert researchers to important aspects of assessment reform, and the comparison of teachers' and researchers' perspectives will help both groups think more effectively

[1]Principals make an important contribution to classroom reform, and the absence from this paper of the principal's perspective is an unfortunate omission forced on us by time and resources limitations.

about these issues. We also believe this framework will help us in subsequent analyses of results from our research in Kentucky and other states.

**Data Sources**

The report is an attempt to conceptualize the teachers' and researchers' perspectives on assessment reform based primarily on information gathered during the development and initial data collection stages of the study. Our insights into important features of state assessments and the way they are viewed by teachers and researchers are based on several sources of information.

- Past research on state assessment systems in California, Kentucky, Maryland, and Vermont and research on local assessment reform efforts in Colorado conducted by the authors and our colleagues contributes to our understanding of the implementation and effects of assessment reform at the school and classroom levels (Borko, et al., 1994; Borko, et al., 1997; Gearhart and Wolf, 1994; Gearhart, et al., 1995; Gearhart and Wolf, in press; Klein, McCaffrey, Stecher and Koretz, 1995; Koretz, Stecher, Klein and McCaffrey, 1994; Koretz, Mitchell, Barron and Stecher, 1996; Shepard, et al., 1996; Shepard and Bliem, 1995; Stecher and Mitchell, 1995; Wolf and Gearhart, 1994; Wolf, Davinroy and Gearhart, in preparation).

- A review of research on the measurement of classroom practices informs our thinking regarding researchers' interests in classroom practices (Consortium on Chicago School Research, 1996; Leighton, et al., 1995; Porter, 1995; Smithson and Porter, 1994; Talbert, 1996).

- Interviews with state educational administrators involved in the development of the Kentucky state assessment system provide insight into desired implementation scenarios at the classroom level.

- Conversations with regional writing resource teachers, PRISM math teachers, and classroom teachers during the development and pilot test of the survey instruments contribute to our understanding of the teachers' perspective on the Kentucky assessment system.

- Interviews with teachers and principals, and observations of writing and mathematics instruction in the case study component of the project[2] provide additional insights into teachers' questions regarding the state assessment system.

**Framework**

On the basis of our interviews, observations and research experience, we developed a framework for studying state assessment reform at the classroom level. The framework includes four broad categories for describing important features of assessment reform: the elements of the assessment, the state and district context, the classroom context, and the attitudes of teachers. These categories subsume eleven specific features of assessment reform that are salient to classroom teachers and to researchers.

Although the same eleven features are relevant to teachers and researchers, the two groups have different perspectives on assessment reform which are also captured in the framework. Teachers typically think in terms of their specific instructional responsibilities; researchers typically think in more general or cross-cutting terms.[3] For example, the two groups agree that stability is important at the

[2]This component of the project focuses on sites that are exemplary with respect to the state reform efforts. In Kentucky, the sample includes three elementary schools and three middle schools. Respondents include one teacher responsible for writing and one teacher responsible for mathematics in the accountability grade levels at each school (grades 4 and 7 for writing; grades 5 and 8 for mathematics).

[3]Research suggests that teachers' concerns about educational innovations evolve over time (Hall, 1979). Initially, teachers' concerns are informational and personal. As teachers become more familiar with the reform, they shift their focus to the management of the change process and its effects within their classrooms. Ultimately, teachers broaden their sphere of concern to include their teaching colleagues in the school as well as other reform options. This evolutionary perspective makes sense to us, but it is not included explicitly in our framework. Rather, we focus on the differences between teachers' and researchers' concerns. In time we may be able to incorporate more about the growth over time in each group's perspectives on assessment reform.

local level, but they differ in how they think about stability. Teachers' comments about the stability of the assessment (e.g., its content, format and rules) relate to the demands that stability or lack of stability places on their time and their willingness to make the effort needed to re-direct their actions. Researchers tend to focus on stability in terms of the genesis of changes that might affect teachers and the pattern of teacher reactions more broadly, e.g., recognition of limitations in the system, political fluctuations, etc.

We find it helpful to describe the teachers' and researchers' viewpoints in terms of the questions they might ask about a particular feature, so the framework is presented in terms of questions. The use of questions is both a rhetorical device and a reflection of the roles that teachers and researchers play with respect to the assessment system. The role of teachers is to discover what the system means and how it will affect their professional lives. The role of researchers is to frame and investigate questions. Hence, the use of questions seems to be a natural way to capture the perspectives of both groups.

Teachers typically ask four types of questions about most aspects of the assessment reform. What is it? Do I agree with it? How does it affect me, my students, and my classroom? What changes do I need to make as a result? Similarly, researchers tend to ask four types of questions when they look at the local effects of assessment reform. What is it? How do teachers feel about it? What effects does it have on practice? What impact does it have on student learning? In addition, researchers are interested in investigating unintended consequences of policies regarding assessment reform.

The framework is presented in Table 1. Although teachers and research ask many questions about assessment, as a simplification the framework focuses on the questions that seem most important to each group. The first column lists the broad categories of assessment features. The second column contains the eleven specific features we identified as salient to assessment reform at the classroom level. The teachers' point of view is captured by questions in the third column, and the researchers' perspective underlies the questions listed in the fourth column.

Following Table 1, each feature is discussed in greater detail. These discussions expand on the key questions and examine the range of concerns one will find within the teacher and researcher communities. The reader should understand that the framework as presented in Table 1 is a heuristic device to simplify a diverse and

complex situation. We do not mean to suggest that there is only one teacher perspective or one researcher perspective. In the discussion of each feature we have tried to represent a range of teacher and researcher interests regarding reform-oriented assessment at the local level. As we learn more about the views of Kentucky teachers, we hope to be able to describe more fully the range of teachers' and researchers' concerns.

**Purpose**

*How do the purposes of the assessment relate to my instructional goals and the needs of my students?* Although teachers are interested in the broader purposes of the assessment (e.g., improving student writing, measuring problem solving ability), their primary concern is how well these purposes match their own instructional goals and objectives. Consequently, teachers' questions about the purpose of the assessment typically focus on understanding how the desired outcomes relate to their current instructional program. For example, they want to know how to interpret the goals of the assessment with respect to the content and format of their lessons and the performance of their students. More specific concerns of teachers include:

- How are the desired outcomes similar to the instructional goals I have for my students, and how are they different?

- Will my teaching methods improve students' ability to achieve these outcomes?

*For what purposes will assessment information be used?* Researchers typically focus their attention on the intended uses of the assessment system. They want to know whether the assessment is designed primarily for monitoring student performance, insuring accountability, or as a lever for change. Subsequent questions reflect these distinctions. For example, in the case of an assessment with a strong emphasis on researchers may want to explore the assumptions implicit in the assessment about what motivates teachers and students. Alternatively, when the assessment system is designed to promote instructional change, researchers may examine the underlying theories of learning on which the assessment is based to see whether these theories are consistent with the beliefs of the teachers who must enact the assessment.

Table 1

Salient Features Of State Assessment Systems At The Classroom Level: Teacher And Researcher Points Of View

| Categories | Features | Teachers' Questions | Researchers' Questions |
|---|---|---|---|
| ELEMENTS OF THE ASSESSMENT | Purposes | How do the purposes of the assessment relate to my instructional goals and the needs of my students? | For what purposes will assessment information be used? |
| | Content | Which aspects of the curriculum does the assessment emphasize and how does it address them? | Which content domains does the assessment cover, and how to teachers match their curriculum to the assessment? |
| | Logistics | What are my students and I required to do, and how great a burden will this place on us? | Who is responsible for which functions, and how will assessment demands be met? |
| | Inclusion and Accommoda-tion | Which students are included, and what accommodations are appropriate? | How consistently are accommodations applied, and what consequences occur as a result? |
| | Accountability | How will the results affect my students, my school, and me? | How do rewards and sanctions affect classroom practice and student learning? |
| | Meaningful-ness | Do the results reflect my students' skills and abilities? | Are the results valid for the intended purposes? |
| STATE OR DISTRICT CONTEXT | Stability | Will the assessment continue in its present form? | How does the historical and political context shape the assessment, past and future? |
| | Training and Support | How do I learn about the assessment and obtain support that is efficient and effective? | Is training and support adequate to enable teachers to implement the assessment and achieve the goals of the reform? |

Table 1, cont.
Salient Features Of State Assessment Systems At The Classroom Level:
Teacher And Researcher Points Of View

| Categories | Features | Teachers' Questions | Researchers' Questions |
|---|---|---|---|
| CLASSROOM CONTEXT | Classroom Practices | How does the assessment fit with my existing teaching and assessment practices? | How does the assessment relate to changes in instructional and assessment practices? |
| | Performance Standards | Do the expectations seem reasonable for my students? | What impact do performance standards have on teachers' expectations for student work? |
| TEACHERS' ATTITUDES | Commitment | How do I feel about the time, energy, and instructional changes that these tests demand? | What factors affect teachers' attitudes toward the assessment? |

Researchers' concerns include:

- What philosophy about the role of assessment in school reform guides the system?

- What is the relationship between teachers' and policymakers' understandings of the purposes of the system?

**Content**

*Which aspects of the curriculum does the assessment emphasize and how does it address them?* Teachers are concerned about the content focus of the assessment and the manner in which knowledge and skills are assessed. For example, Kentucky teachers want to know which aspects of curriculum are covered in KIRIS, the way they are tested, and the relative importance assigned to each. If the assessment includes topics not present in the classroom curriculum, teachers must decide whether and how to change their teaching emphasis to accommodate the new content, including which topics to reduce or eliminate. If the assessment measures things in new ways (e.g., collaborative activities, portfolios), teachers will need to figure out how to prepare students to demonstrate their knowledge and skills in the

new formats. Information about the emphasis of the test permits teachers to assign priority to those topics that receive priority on the assessment.

Complications can arise for teachers if there are other policies in place that emphasize different aspects of the curriculum than the state assessment. One tenet of the current movement toward "systemic" reforms is that curriculum, assessment, instructional materials, and staff development should all be coordinated. However, in states where assessment reform is used to drive instructional change, such "alignment" may not happen for quite some time. Even in states where curriculum reform and assessment reform go hand in hand, there can be a delay of years between the implementation of the assessment system and the adoption of new textbooks and instructional materials that are consistent with the emphasis of reforms. More specific concerns of teachers include:

- Am I teaching the topics that will appear on the test?
- How much emphasis should I give to the various topics I am teaching?
- How does the assessment system align with my textbooks and other instructional materials?
- How do I find classroom time to include the material covered in the assessment?

*Which content domains does the assessment cover, and what type of measures are included?* Researchers are concerned with the relationship between the content domains sampled in the test and the content emphases implicit in teachers' classroom curriculum, but their interest focuses on the process of change. They want to know what teachers think the assessment is designed to measure, how teachers' understandings compare with test developers,' how the test's content compares with the classroom curriculum, and how teachers adapt their curriculum to the assessment. Researchers may be particularly interested in how teachers try to address more complex skill domains, such as problem solving or higher-order thinking, that may be covered in the assessment. Researchers also are interested in factors that affect teachers' efforts to make changes consistent with the content and methods of the test. Researchers' specific concerns include:

- What is the match between the teacher's classroom curriculum and the test content?

- Which domains do teachers think are emphasized by the assessment?

- Do teachers change the content of the classroom curriculum to match the test content? If so, how?

- Which topics do teachers have the greatest difficulty addressing?

- Which factors determine teachers' success in addressing particular content areas?

**Logistics**

*What are my students and I required to do, and how great a burden will this place on us?* Teachers are concerned about new demands the assessment places on their classroom and preparation time. If an assessment is complex, then teachers will have a great deal of concern about logistics. Details, such as how the assessment is to be integrated into the school year, what responsibilities teachers will have for administering the assessments, and how they will be expected to prepare students to complete the activities, will be important.

Teachers' concerns are likely to be heightened if the assessment contains more than on-demand, multiple-choice tasks. Assessments that embody constructed-response questions and cumulative performances (e.g., portfolios) raise additional questions. For example, portfolios require teachers to change instructional plans and re-allocate class time on an ongoing basis (not merely once a year during a brief assessment period). Performance assessments also raise questions for teachers about the types of support and guidance they should provide to their students.

Logistical concerns have tremendous importance to teachers because they are constantly making trade-offs between competing uses for limited school time. For example, the time taken to administer the assessment is time taken away from other things, most notably instruction. Teachers often perceive state assessments to be "add-ons" that must be completed *in addition to* their ongoing responsibilities rather

than as supplemental activities that can be *incorporated into, and add value to,* their instructional programs. As a result, logistics can have a significant impact on teachers' support for the assessment. In contrast, some teachers we talked with see the new reform as a validation of work they have been doing for years. Rather than an add-on, they well understand how to integrate the reforms into their existing programs. Still, teachers' support is generally inversely related to the amount of time required to administer the assessment and the degree to which its perceived burden exceeds its perceived value.

Teachers' specific concerns include:

- When will the assessment be administered?
- What will I need to do to prepare for the administration (e.g., organize booklets or materials)?
- What responsibilities will I have during the test administration?
- How much time will the test administration involve?
- Where will students take the test?
- What components will be administered in my class?
- What components must I prepare my students for?

*Who is responsible for which functions, and how will assessment demands be met?* Researchers want to understand the allocation of responsibility for the assessment and how teachers fit the new demands into their instructional programs. Many of these demands appear to be quite simple, and researchers may be tempted to underestimate the importance of assessment logistics to teachers. However, researchers should try to understand the trade-offs teachers face when implementing a new assessment systems and the importance of administrative demands in their planning.

**Inclusion and Accommodation**

*Which students are included in the assessment, and what accommodations are appropriate?* Teachers' primary concerns about the way the assessment treats students who have special needs are practical: Who should be tested (inclusion)? What assistance should be given during the assessment (accommodation)? Teachers need to know the assessment system's rules concerning special needs students. They also need to know how to apply the rules for particular students, a question which is often complicated because it is subject to interpretation.

The situation in Kentucky is relatively clear. The assessment reform emphasizes the inclusion of all students regardless of disabilities. One to two percent of all students with moderate to severe cognitive disabilities are eligible for an alternative assessment program, called the Alternate Portfolio, but the overwhelming majority of student participate in the regular KIRIS assessment. Students with IEPs or Section 504 plans may be tested with either accommodations—changes in the testing environment or process—or modifications—changes in the testing instrument. Furthermore, Kentucky specifies that such accommodations or modification are only permitted if they are explicitly included in the student's IEP or Section 504 plan and are regularly provided to the student as part of his or her instructional program. (They cannot be introduced for the purposes of the assessment.) Accommodation and modifications include paraphrasing, oral presentation of the assessment, allowing dictation of responses, cueing, use of an interpreter, and technological aids (Koretz, 1997).

Kentucky provides detailed guidelines about these alternations. These guidelines clarify many of the questions teachers might have about how to implement them, but they still leave areas for interpretation. Furthermore, not all states are as explicit as Kentucky. In many cases teachers are uncertain about how they should treat students during the assessment, e.g., whether they are permitted to provide a reader for a dyslexic student or extra time for a learning-disabled student.

Teachers also need to know how the performance of their special-needs students will affect the overall distribution of scores. If teachers are to be held accountable for their students' performance, they want to know how the results will take into consideration the special needs and backgrounds of the students in their classes. Specific teacher concerns include:

- Which students are eligible for accommodations?
- What type of accommodations are appropriate?
- Can I use multiple accommodations for the same student?
- What assistance is available if I am uncertain about including particular students or providing accommodations?
- Will the rules regarding special needs students place extra demands on my time?

*How consistently are accommodations applied and what consequence occur as a result?* Researchers also want to understand the rules regarding inclusion and accommodations. They are concerned about how consistently these rules are applied and the effects these policies have on regular classroom practices. States have different rules about which students are required to participate in testing programs and under what conditions tests are administered. In the past, many Special Education students and students whose native language was not English were excluded from statewide testing. Increasing numbers of states are changing their rules to reduce or eliminate these exclusions to promote the goal of having all students participate in the assessment in ways that are appropriate.

Researchers are interested in measuring the consistency with which assessment exclusions and accommodations are applied. Are there differences within schools, districts or regions? Do students with certain disabilities receive more consistent treatment than students with other disabilities? Researchers are also interested in the effects of assessment accommodations on other instructional activities. For example, does the policy connecting accommodations to IEPs affect the number of students with IEPs or the types of accommodations that are included in students' IEPs?

**Accountability**

*How will the results affect my students, my school, and me?* Teachers are extremely concerned about the use of assessment results for accountability purposes, and they take any accountability provisions of state assessments very seriously and very

personally. It does not seem to matter whether there are monetary rewards associated with improvement, such as those used in Kentucky, or non-monetary stakes in the form of public reporting of school comparisons, such as those that occur in Maryland; teachers seem to care about formal accountability in either case. They want to know what kinds of decisions will be based on test scores and how these decisions will affect their students, their school, and themselves.

Teachers want to know who will be held accountable. Many testing programs include only three or four grade levels, so not all teachers' students participate in the assessment. We noticed differences among Kentucky schools in the informal assignment of responsibility for the assessment results. Some Kentucky schools create an atmosphere in which the responsibility for student performance belongs to the entire school, not just teachers in the KIRIS grade levels. At other schools, the focus is on teachers in the accountability grades. For example, assignments of teachers to grade levels were switched in some Kentucky schools to match changes in the accountability grades from eighth to seventh in writing and from fourth to fifth in math. Moreover, schools differ in the seriousness with which they view the assessment results. The names of the top-performing schools were posted on the principal's door in one school we visited. Posters with KIRIS performance levels and definitions were displayed throughout the school. In this case, the school had taken accountability to heart; faculty also were clear that the responsibility for performance belonged to the school as a whole, not to a select set of teachers.

Specific questions teachers might ask include:

- Do the scores on the assessment reflect what my students know and can do?
- Are the score from one class comparable to the scores from another?
- What classroom activities will improve student scores?

*How do rewards and sanctions affect classroom practice and student learning?* Researchers think about the accountability aspects of assessment reform in more general terms. They want to know whether the accountability mechanism has the desired result of promoting improved student performance and how the elements of the accountability system affect principals' and teachers' behaviors. For example, they are interested in the relationship between teachers' attitudes toward

accountability and student progress. Researchers want to know if there is a difference in outcomes between schools that see assessment results as a shared responsibility of all teachers and schools that see results as the responsibility of individual teachers. Similarly, they are curious to know whether there are differences in results between schools where the strongest teachers were moved to accountability grades and schools where this did not occur.

Other specific concerns of researchers include:

- How do teachers' attitudes toward accountability relate to school and student performance?

- To what extent do rewards and sanctions motivate schools to change?

**Meaningfulness**

*Do the results reflect my students' skills and abilities?* Teachers are concerned that the assessment results provide a fair indication of the skills and abilities of their students. This is particularly true when the assessment system is used for the purpose of accountability (i.e., when high stakes are attached to performance). Although teachers may not use formal measurement terminology, their questions about the accuracy or meaning of student scores are questions about the validity of the results. Such concerns may be triggered if teachers notice differences between students' performance in class and their performance on the assessment or differences between the format of the test and the way material is taught in class. Similarly, teachers may comment about the amount of writing required to complete mathematics portfolios or about differences among teachers in the number of times students are allowed to revise work in their writing portfolios. These, too, are concerns about the validity of scores.

Specific questions teachers might ask include:

- Do the scores on the assessment really indicate what my students know and can do?

- Do all students have the same chance to score well on the assessment?

- What can I do to improve students' performance on the assessment?

*Are the results valid for the intended purposes?* Researchers are also concerned about the reasonableness of scores, but they have a more formal vocabulary for discussing it. This vocabulary is based on a rich theory of educational measurement that helps researchers conceptualize the factors that affect student scores. At the heart of researchers' concerns about meaningfulness is the question of validity—are the scores appropriate for the intended use? Current assessment reforms place greater emphasis on constructed-response tasks and performance events. These types of activities have the potential to provide valid measures of many important skills and ability. However, they also are more difficult to develop, administer, score and interpret.

A wide variety of factors affect the validity of scores, including the design of the assessment, the administrative procedures, the scoring rules, and the purposes for which the scores will be used. A recent study of the KIRIS by an independent panel of measurement psychologists identified a number of problems that reduced the validity of scores for accountability purposes, and the Kentucky Department of Education has reacted swiftly to address them (Hambleton, et al., 1995).

Many of the elements of the framework that have been described above have implications for validity. For example, researchers may raise questions about the relationship between the goals of the assessment and the format that is being used to measure student performance (e.g., on-demand measures versus cumulative measures, selected-response versus constructed-response). Similarly, when researchers think about logistics, they are concerned that the conditions under which the assessment is administered may affect the results. If the assessment is administered in a classroom where educational material is displayed on the walls, this could affect students' scores. Researchers raise additional concerns about the logistics of cumulative assessments, such as portfolios, where teachers play a significant role guiding students' work. When teachers control the environment in which assessment pieces are created, it is important to ensure consistency in the support they provide students when students are preparing their work.

Other specific concerns of researchers include:

- Are the measures appropriate to the constructs being measured?

- Are performance assessments scored reliably?
- Are the results valid for the purposes of accountability?
- Are the measures fair for all groups of students?
- Are the results credible in the eyes of the public and the professional community?

**Stability**

*Will the assessment continue in its present form?* Meeting the goals of assessment reform often requires major changes in what is taught and how it is taught, and teachers may not be willing to make these changes unless they believe the reform will last. For example, many new assessments involve writing across the curriculum, but teachers in mathematics and other subjects traditionally do not emphasize writing. Many of these teachers have had little training in teaching writing. To incorporate writing into their classroom practices, they must learn new skills, develop new instructional activities, and learn to evaluate students in new ways. Before they make the investment in new skills and activities, teachers often want to know whether the assessment program is likely to change rapidly.

Experienced teachers sometimes comment that reform efforts come and go, in some cases quite quickly. Further, even when reforms are here to stay, they sometimes undergo substantial modification. For example, Kentucky changed accountability grades; the writing assessment moved from grade eight to grade seven, and the mathematics assessment moved from grade four to grade five. In addition, the mathematics portfolio component was put on hiatus for a year while the conceptual framework for scoring was rethought. Some change can be understood as improvement. Too much change is often interpreted as instability, capriciousness, or whim. Further, regardless of how they interpret change, teachers can lose their willingness to work to modify their instructional practices if they perceive that the system is likely to change in substantial ways. Thus, with an assessment-based reform, as with any other reform, teachers want to know whether the system is likely to remain relatively constant or whether it is changing frequently in terms of subject, grade levels, components, scoring procedures, etc. In the latter

case they may ask themselves, "If I hunker down and ignore this, will it eventually go away?" Other teachers seem to indicate just the opposite stance. They feel that the assessment reform is here to stay and are glad to weather the sudden shifts and modifications that might be necessary to achieve the larger goal of authentic and appropriate reform.

*How does the historical and political context shape the assessment, past and future?* Researchers often approach the question of stability in terms of external factors. They want to understand the historical and political sources of stability and instability in the assessment reform. These sources might include the state's financial commitment to the reform, the degree of bipartisan political support for the reform, and the presence of vocal constituencies that support or oppose the reform (McDonnell, 1997).

Researchers are also interested in teachers' perceptions of stability of the reform, and how these perceptions shape teachers' actions. They want to know if there are differences in teachers' judgments about the stability of the reform and what factors determine these differences. Researchers also want to know about relationships between teachers' perceptions and their willingness to make changes. Researchers' concerns include:

- How broad is the political base of support for the assessment system?
- Is the assessment system associated with one political or administrative faction or another?
- Does assessment have its own line in the budget or is it subject to annual reauthorization?
- How does the assessment system fit in terms of the history of educational changes in state, particularly changes in curriculum?

**Training and Support**

*How do I learn about the assessment and obtain support that is efficient and effective?* When a new assessment system is implemented, the teachers' initial concern is to

understand what the system entails: the rules, procedures, and guidelines that must be followed. After these questions are answered, teachers start thinking about the implications of the assessment for their classroom practices. If they want additional information, they begin to look for relevant professional opportunities. In Kentucky, teachers have a wide range of opportunities to learn more about KIRIS and what they can do to promote desired student outcomes. The opportunities include regional workshops, university-based professional development, and local in-service training, as well as instructional television programs and opportunities for classroom observations.

However, teachers are usually under competing pressures for their time, and they are looking for training that is efficient. They want to learn about the assessment system, but they do not want to take time away from class to attend institutes during the day, nor do they want to give up their personal time. They are looking for assurances that the professional development activities they participate in will be worth the time they invest.

Many teachers are also concerned about the effectiveness of professional development. Some teachers prefer being shown rather than told what to do. They would like to see what good instruction looks like in the classroom and prefer observing other classrooms or having trainers come to their classroom and demonstrate for them. Other teachers prefer it when trainers observe in their classroom and provide feedback based on these observations. Some teachers are also interested in becoming involved in training as providers of services beyond their classroom responsibilities. In Kentucky, teachers can explore opportunities such as becoming cluster leaders or portfolio scorers.

Specific teacher concerns include:

- What workshops are available that will help me prepare my students for the assessment?

- Will this professional development opportunity provide the information I want in a reasonable amount of time?

- What support is provided to me if I enroll?

- Are the workshops offered at convenient times and locations?

- Who can I rely on locally to respond to my questions and provide ideas to help me prepare my students for the assessment?

*Is training and support adequate to enable teachers to implement the assessment and achieve the goals of the reform?* Researchers are gaining a better appreciation of the role of professional development in assessment-driven reform. If the goal is to change classroom practice, it is often not enough to simply require students to take an assessment that emphasizes important content. Teachers must also be shown how to teach in new ways that promote learning of this new content. Researchers think about how support is organized and delivered and who has access to various professional development activities.

Researchers are also interested in the choices teachers make regarding professional development and differences between teachers, schools, and districts in the use of various training and support options. For example, sites identified as "exemplary" in terms of instructional practices seem to have more teachers who are themselves trainers, cluster leaders, or in-service developers. Finally, researchers are interested in the relationship between professional development and assessment results. Specific researcher concerns include:

- How do teachers learn about the assessment?

- Is professional development an integral part of the assessment system?

- What formal mechanisms are created for training teachers to understand and implement the new system?

- At what level are these services provided—local, district, regional or state—and under whose auspices?

- How much professional development do teachers participate in?

- How does the assessment system encourage informal networks and communities of practice?

- Is participation in professional development related to assessment results?

**Classroom Practices**

*How does the assessment fit with my existing teaching and assessment practices?* Teachers' concerns about content (discussed previously) are closely related to their concerns about teaching and testing practices. Not only must teachers decide how to adapt their curriculum to the content emphasis of the tests, they must decide whether to change their instructional strategies and their classroom assessment activities, as well.

Our conversations with teachers suggest that their endorsement of the assessment system is related, in part, to the ease with which the new assessments can be incorporated into ongoing instructional and assessment practices. In particular, teachers may react negatively if the assessment adds new responsibilities—with respect to content, instructional practices, or assessment—without permitting teachers to reduce existing demands. For example, Kentucky teachers may want to spend more time having students write about mathematics to prepare them for the mathematics portfolio assessment, but they may also feel pressures from parents to continue their mathematics drills to maintain students' basic computational skills. Similarly, some districts we have visited as part of other CRESST research had their own testing requirements, such as timed tests of math facts. Although teachers were willing to incorporate performance assessments into their instructional programs, they would not do so at the expense of the timed tests. For teachers in Kentucky who are using chapter tests in their mathematics texts, portfolios may be perceived as placing additional—and not necessarily compatible—demands on their assessment practices. Demands that are perceived as "add-ons" rather than substitutions can be particularly frustrating. In contrast, some teachers support the assessment because it is consistent with changes they want to make in their classroom but were unable to make due to lack of support from school administrators. In essence, the assessment system gives their views more clout. Teachers' specific concerns include:

- Is my current teaching style appropriate to the new assessment?
- Do I need to change the way I present material to accommodate the (performance) measures used on the assessment?
- How does the assessment system relate to my existing classroom assessment methods and purposes?
- Is the assessment compatible with existing district assessment requirements?

*How does the assessment relate to changes in instructional and assessment practices?* Researchers are also concerned about the match between the demands of the assessment reform and current instructional and assessment practices, but their emphasis is on understanding patterns in teachers' responses to any perceived differences. Researchers are interested in why teachers do or do not change practice, and whether there are identifiable factors that predict changes. For example, in a previous CRESST study, Borko and colleagues found that teachers did not change their practices much at all if the distance between the assessment reform and current practice was too great. This suggests small, incremental changes may be easier to achieve than large, decisive changes. Researchers' concerns include:

- How much do teachers change their teaching and testing practices?
- What changes are most and least common?
- What differences are associated with teachers who make changes and teachers who do not?
- What other assessments are in use, and how do these change?

**Performance Standards**

*Do the expectations seem reasonable for my students?* Teachers typically voice a strong desire to understand what students are expected to know and to what level they must perform. Teachers we talked with placed considerable emphasis on the

need to understand performance criteria and decide whether they were appropriate for their students. However, our experience suggests that there are considerable differences among teachers' beliefs about student capability. Some teachers hold high expectations for all students and do not find external performance standards to be a problem. Others are concerned that uniform expectations are not reasonable for all students. They endorse the belief that all children can learn, but not always that all children can learn to the same high levels. These differences may affect how teachers think about the reasonableness of the expectations that are implicit in the assessment system.

Teachers are also interested in the impact of performance standards on students' motivation. Certainly students react differently to high standards, and teachers pay careful attention to students' responses, looking for ways to utilize the assessment standards effectively as motivational tools. Teachers' specific concerns include:

- Does the assessment system focus on outcomes I endorse?
- Does it set standards or expectations that are appropriate for my students?
- Are the goals for learning gains reasonable for my students?
- When do standards encourage students and when do they place too much pressure on students?

*What impact do performance standards have on teachers' expectations for student work?* Researchers are interested in whether the high expectations underlying performance standards affect teachers' expectations for their students. For example, assessments that incorporate standards for students' higher order thinking skills and problem solving capabilities may lead teachers to raise their expectations in these areas. Researchers are also interested in whether the impact on teachers' expectations differs for different groups of students. Researchers' specific concerns include:

- Do teachers agree with the performance standards?

- Do teachers who are involved in setting performance standards change their expectations more than teachers who were not involved?

- What impact do performance standards have on student motivation to perform well on the assessment?

- What impact do performance standards have on student achievement more broadly?

**Commitment**

*How do I feel about the time, energy, and instructional changes that these tests demand?* Teachers want to know whether the benefits of the new assessment—for their students and themselves—are worth the effort they must expend to implement it. If so, most teachers will bear the burdens willingly; if not, many will be reluctant to make the effort. Assessment is not unique in this respect; teachers make similar decisions daily when they decide how to use their limited classroom time and resources most effectively. A new assessment represents another competing demand that must be factored into their continuous calculations.

Some teachers undoubtedly find it easier to resolve these demands in favor of the status quo. They choose to avoid the new burdens that assessment reform places on them. In contrast, others welcome the opportunity for change that the new assessment creates. Teachers' level of commitment depends on many factors, including how the assessment is presented to them, their belief in the goals of the assessment, the logistical demands of the assessment, their perception of the longevity of the assessment, and the consequences of their students' performance of the assessment. Teachers' specific concerns include:

- Do I believe that the assessment measures important outcomes?

- If I make an effort, will my students' scores improve?

- What is the likely relationship between the amount of effort I make and changes in my students' scores?

- Do I know anyone who is comfortably working with the assessment?

- Can I picture how these changes would look in my classroom?

*What factors affect teachers' attitudes toward the assessment?* Researchers recognize that teachers must be fully engaged in implementing the assessment for the goals of the assessment to be achieved. They are interested in understanding the factors that affect teachers' commitment to the assessment and the relationship between teachers' level of commitment and student scores. Test scores may be related to teacher commitment in unexpected ways. For example, teachers who take on the challenge of making the fundamental instructional changes embodied in an assessment reform may find that their students' scores remain constant or decline for a year or two because of the difficulty of making these changes. In contrast, teachers who focus more narrowly on the content of the test may produce high initial gains in student scores. Thus, researchers are interested in teacher commitment as a contributing factor to student outcomes. Researchers' specific concerns include:

- How much variation is there in commitment to the assessment within a school and between schools?

- What factors explain differences in teachers' commitment to assessment reform?

- How do principals and teachers interpret annual changes in assessment results, and how do these changes affect their decisions regarding practices?

**Summary**

This framework captures a number of features that are important in understanding the effects of assessment systems at the local level. Broadly speaking, local effects are related to the ways in which teachers come to understand the principles that guide the assessment system, integrate these principles with existing policies and procedures, adapt existing practice to promote desired outcomes, and do this within the context of their local operating environment. The framework

highlights eleven interrelated features that play important roles in determining how assessment reforms will affect classroom practices. There probably are others we have not yet identified.

The framework also compares the perspectives of researchers and teachers about these features. Researchers who hope to contribute to improvements in practice must attempt to understand the local impact of assessment reforms from the perspective of the teachers first, and then look for patterns of responses across teachers and classrooms and trends in responses over time. Researchers who ignore the individual teacher's perspective may overlook factors that are important in shaping local responses to assessment reforms. Following our own advice, we have designed our study of the effects of state assessment reforms on classroom practices to begin with the perspectives of individual teachers. We hope that by doing so, we will be able to clarify the relative importance of these and other features of assessment reform to local practice and to recommend ways to foster practices that improve student performance.

## References


Aschbacher, P. R. (1991). *Alternative assessment: State activity, interest and concerns.* CSE Technical Report 322. Los Angeles: UCLA Center for Research on Evaluation, Standards and Student Testing.

Bond, L. A., Braskamp, D., van der Ploeg, A., & Roeber, E. (1995). *State student assessment programs database: School year 1994-1995*. Oak Brook, IL: North Central Regional Educational Laboratory.

Borko, H., Davinroy, K. H., Flory, M. D., & Hiebert, E. H. (1994). Teachers' knowledge and beliefs about summary as a component of reading. In R. Garner & P.A. Alexander (Eds.) *Beliefs about text and instruction with text* (pp. 155-182). Hillsdale, NJ: LEA.

Borko, H., Mayfield, V., Marion, S., Flexer, R, & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Success, stumbling blocks, and implications for professional development. *Teaching and Teacher Education*, 13(3), 259-278.

Consortium on Chicago School Research. (1996, September). *The Charting Reform survey series: Public Use Data.* Version 1.1. Chicago: Author.

Gearhart, M., Herman, J. L., Novak, J. R., & Wolf, S. A. (1995). Toward the instructional utility of large-scale writing assessment: Validation of a new narrative rubric. *Assessing Writing*, 2(2), 207-242.

Gearhart, M., & Wolf, S. A. (in press). Issues in portfolio assessment: Assessing writing processes from their products. *Educational Assessment*.

Gearhart, M., & Wolf, S. A. (1994). Engaging teachers in assessment of their students' narrative writing: The role of subject matter knowledge. *Assessing Writing*, 1, 67-90.

Hall, G. E. (1979). The concerns-based approach to facilitating change. *Educational Horizons*, 57(4), 202-208.

Hambleton, R. K., Jaeger, R. M., Koretz, D., Linn, R. L., Millman, J., & Phillips, S. E. (1995). *Review of the measurement quality of the Kentucky Instructional Results Information System, 1991-1994. Final Report*. Frankfort: Office of Educational Accountability.

Klein, S. P., McCaffrey, D., Stecher, B., & Koretz, D. (1995). The reliability of mathematics portfolio scores: Lessons from the Vermont experience. *Applied Measurement in Education*, 8(3), 243-260.

Koretz, D. (1997). *The assessment of students with disabilities in Kentucky*. *Draft Deliverable*. Los Angeles, CA: CRESST/UCLA.

Koretz, D., Mitchell, K, Barron, S., & Stecher, B. (1996). *Perceived effects of the Kentucky Instructional Results Information System (KIRIS)*. RAND, MR-792-PCT/FF.

Koretz, D., Stecher, B. M., Klein, S., & McCaffrey, D. (1994, Fall). The Vermont portfolio assessment program: Findings and implications, *Educational Measurement: Issues and Practices,* 13(3), 5–16. (Reprinted as RAND, RP-366, 1995.)

Leighton, M. S., Mullens, J. E., Turnbull, B., Weiner, L. K., & Williams, A. S. (1995, March). *Measuring instruction, curriculum content, and instructional resources: The status of recent work*. Working Paper No. 95-11. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

McDonnell, L. M. (1997). *The politics of state testing: Implementing new student assessments.* CSE Technical Report 424. Los Angeles, CA: CRESST/UCLA.

Porter, A. C. (1995, September). Developing opportunity-to-learn indicators of the content of instruction: Progress report. Madison: Wisconsin Center for Education Research.

Shepard, L. A., & Bliem, C. L. (1995). Parents' thinking about standardized tests and performance assessments. *Educational Researcher*, 24(8), 25.

Shepard, L. A., Flexer, R. J., Hiebert, E. H., Marion, S. F., Mayfield, V., & Weston, T. J. (1996). Effects of introducing classroom performance assessments on student learning. *Educational Measurement: Issues and Practice,* 15(3): 7-18.

Smithson, J. L., & Porter, A. C. (1994, October). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum-—The reform up close study*. Consortium for Policy Research in Education Report #31. New Brunswick: Rutgers, The State University of New Jersey.

Stecher, B. M., & Mitchell, K. J. (1995). *Portfolio driven reform; Vermont teachers' understanding of mathematical problem solving,* CSE Technical Report 400, Los Angeles, CA: CRESST/UCLA.

Talbert, J. (1996). *Students-at-the-center teacher survey*: *Elementary grades questionnaire: Philadelphia.* Stanford: Center for Research on the Context of Teaching.

Wolf, S. A., Davinroy, K., & Gearhart, M. (in preparation). "The clay that makes the pot:" The loss of language in writing assessment.

Wolf, S. A., & Gearhart, M. (in press). New writing assessments: The challenge of changing teachers' beliefs about their students as writers. *Theory into Practice.*

Wolf, S. A., & Gearhart, M. (1994). Writing what you read: Assessment as a learning event. *Language Arts*, 71(6), 425-444.