Teacher views of students in the gaps

Caroline E. Parker, Education Development Center, Inc.
Susan Saxon, Educational Alliance at Brown University

January 2007

Support for this investigation was funded in part by Grant CFDA #84.368 of the U.S. Department of Education, Office of Elementary and Secondary Education, awarded to the Rhode Island Department of Education on behalf of the New England Compact. The opinions expressed in this paper do not necessarily reflect the position or policy of the Department of Education, and no official endorsement should be inferred.

Table of contents

Introduction 3
Background 4
Methods 5
Findings 6
  Teachers describe two gaps in the assessment system 7
  Students with disabilities are found in both gaps 10
    Short-term memory 10
    Focus 11
  ELLs are found in both gaps 11
  Other student attributes contribute to assessment gaps 12
    Different learning styles 13
    Motivation and/or interest 15
  Vignette: An alternative learning setting 17
  Access to grade-level curriculum 18
  Out-of-school issues can affect assessment results for all learners 19
Conclusion 20
References 22
Appendix A. Interview protocol for initial teacher interviews: Looking at assessment gaps 24

Introduction

The last 10 years have seen rapid changes in the development of large-scale assessments. The No Child Left Behind Act of 2001 (NCLB) requirements that districts and schools measure student achievement, including that of students with disabilities and English language learners (ELLs), have created pressure to improve the ability of large-scale assessments to accurately measure student achievement and ensure access for all students. To this end, assessments using universal design principles promote increased access through various means, such as font size, use of blank space, and use of illustrations in test construction (Thompson, Johnstone, Anderson, & Miller, 2005; Universally Designed Assessments, "Universally Designed Assessments," 2005). States have also increasingly standardized their use of accommodations to increase access (Bolt & Thurlow, 2004; Edgemon, Jablonski, & Lloyd, 2006; Ysseldyke, Thurlow, Bielinski, House, Moody, & Haigh, 2001). Other research has led to the development of new English language proficiency tests focusing on academic English acquisition in four domains (listening, speaking, reading, and writing). States are now required to administer an English language proficiency test to all current ELLs and for two years following reclassification. The results of these new language proficiency tests may help to better understand ELL performance on content assessments (Abedi & Dietel, 2004; August, 2006).

Yet, even with all of these changes and the resulting increase in access for many students, educators argue that the current assessment structure as mandated by NCLB still does not reliably measure the achievement of *all* students (Almond, Quenemoen, Olsen, & Thurlow, 2000; Quenemoen, Thompson, Thurlow, & Olsen, 1999; Wiener, 2006). A significant number of students may not be well-served by the current generation of large-scale assessments inasmuch as the assessments do not reliably measure the achievement of certain groups of students. These gaps in the assessment structure are not well understood, and there is little research on either the gaps or the students who fall into them.

This paper forms part of a larger study of students in the assessment gaps. The study looked at the issue of the assessment gaps from multiple positions and through multiple methods. Teachers were interviewed to better understand their views of large-scale assessment and students in the gaps; large-scale assessment data were analyzed to identify students in assessment gaps; student questionnaires, completed as part of the large-scale assessment, were analyzed to understand some aspects of the student perspective; and multiple assessment modalities were piloted and analyzed. Each paper addresses one of five overarching research questions:

1. Who are the students in the gaps?
2. Of all students who are not proficient, how can states identify those who are in the gaps?
3. What are the attributes of students in the gaps, and how do these students perform?
4. What issues in the assessments themselves contribute to the gaps?
5. Are there specific aspects of the multiple-choice items (complexity, presentation) used in state assessments that contribute to the assessment gaps?

The study described in this paper focuses on the first question—the characteristics of students not served well by current large-scale assessments. This exploratory study looks at the question from the point of view of middle school mathematics and special education teachers. We chose to begin with teachers because they deal with students daily in the classroom, and they are the

ones most likely to identify any discrepancies between the day-to-day classroom-demonstrated proficiency of their students and their large-scale assessment results. A second research paper using the same data addresses the third question of the larger study: *What issues in the assessments themselves contribute to the gaps?*

This study takes place in the context of the purpose of large-scale assessment within NCLB: accountability for schools. If there are students who are working proficiently at grade level, but this is not being captured by the assessment, the schools will be unfairly judged as not meeting adequate yearly progress. If there are students who are working well below grade level but who are progressing in their learning, a grade-level assessment will not show that progress, and again, the schools will be unfairly judged. Thus, the stakes are high for getting assessments right.

Background

Federal legislation over the last 20 years has focused on increasing access to instruction for students with disabilities through the "least restrictive environment." IDEA 2000 legislates that students with disabilities must be included in large-scale assessments for accountability purposes, and NCLB regulations state that schools must look at students with disabilities as a separate subgroup and be accountable for their progress and achievement. Early research on students with disabilities and assessment focused on improving test accessibility through the use of such formats as increased font size and less crowded pages (Bielinski, Thurlow, & Ysseldyke, 2001; Dolan, Murray, & Strangman, 2006). More recently, research has focused on increasing access through accommodations (Dolan et al., 2006), although Tindal et al. found that teachers did not make good decisions about which students would benefit from a read-aloud accommodation in mathematics (Helwig & Tindal, 2003), indicating that simply *providing* accommodations is only the first step in allowing students to take full advantage of them. Research continues to explore the relationship between accommodations and assessment results (Exceptionality, 2004; Shaftel, Yang, Glasnapp, & Poggio, 2005). The "new generation" of assessments and accommodations has been designed to allow more students with disabilities to be accurately assessed, but there is little research looking at whether these types of changes are sufficient for all students with disabilities.

ELLs have faced a similar trajectory with large-scale assessments, though research on the most effective ways to measure these students'content knowledge is still in the early stages. Research on accommodations indicates that the use of sheltered English is the most effective accommodation (Abedi, Courtney, & Leon, 2003; Francis, Rivera, Lesaux, & Rivera, 2006; Hofstetter, 2003; McCardle, Mele-McCarthy, & Leos, 2005). NCLB requires that all ELLs be assessed on content knowledge after one year, and that after three years they must take the regular assessment. They also must be annually assessed on their English proficiency. However, there continues to be a dearth of research on the experiences of ELLs with both the regular assessment and the English language proficiency test.

Both the legislation and the research agendas have recognized the challenge of effectively assessing students with disabilities and ELLs. However, there are other students who also struggle with large-scale assessments. In many states, the absolute number of general education students scoring below proficient is greater than the absolute number of students with disabilities or ELLs. This is an important group of general education students deserving of further research.

For example, do different student learning styles have an impact on student ability to demonstrate proficiency on a large-scale assessment? What is the effect of student motivation on assessment results? There is very little literature in this area, and there is very little research looking at the assessment needs of students who, for whatever reason, are working far below grade level. Looking at the assessment gaps with this lens is a recent phenomenon, and there is very little literature that examines issues of learning styles, motivation, or out-of-school contexts on assessment results.

The literature does not answer the question with which we began this research: Who are the students who are not served well by the current generation of large-scale assessments and thus are in an "assessment gap"? We chose to begin the research with an exploratory, qualitative study focusing on the perspectives of teachers who are in daily contact with students and who also see them in the large-scale assessment context. Our exploratory study sought to more carefully identify those factors that contribute to students not being served well by the large-scale assessments, including but not limited to students with disabilities and ELLs. We sought to describe teachers' understandings of their own students' achievement in eighth grade mathematics, both in the classroom and on the large-scale assessment.

Methods

Twenty-three mathematics teachers, 14 special education teachers, and 3 administrators with special education or mathematics expertise from the four New England states were interviewed, for a total of 40 participants. The majority were eighth grade teachers, except in New Hampshire, where the middle-level 2004 mathematics assessment was administered in sixth grade. Interviews were conducted using an open-ended protocol, and access to specific student scores was provided, when available, as a starting point for discussion.

The interviews for this study were conducted in the spring and early fall of 2005, before the implementation of the NCLB requirement that students be tested annually in grades 3–8. In three of the states, the large-scale assessment has been totally restructured, and thus it should be noted that the interviews refer to the previous generation of assessments.

We chose to conduct our research with eighth grade mathematics for two reasons. We chose mathematics because there has been less research on mathematics assessments than on reading assessments. We chose eighth grade because three of our four states had conducted assessments of that grade level in 2004, when we began the study (we used sixth grade teachers in New Hampshire, where the assessment was in sixth grade).

As an exploratory study, and because of time constraints, we decided that it would be best to choose our sample based on convenience. State education officials provided a list of three or four schools in each state whose teachers and administrators might be willing to participate in the study and that reflected the diversity of schools in their state: large, medium, small, urban, rural. We then contacted the schools directly and set up the interviews. In some cases, the teachers had access to student assessment results, though in others they did not. In each school, all eligible teachers (mathematics or special education teachers at the appropriate grade level) were invited to be interviewed (see Appendix A for the interview protocol). In the smaller schools, all eligible

teachers participated; in the larger schools, a selection of teachers participated. The following table provides an overview of the participating schools.

Table 1. Participating schools

	Urban status	School size	Grade span	% Caucasian	% free/ reduced -price lunch	Other	Teachers		
							Math	SPED	Other
State 1, school 1	Suburban	933	6–8	88	39		3	1	
State 1, school 2	Suburban	630	6–8	93	17		2	1	
State 2, school 1	Rural	579	7–8	98	35		2	2	
State 2, school 2	Rural	388	6–8	97	20		2	2	1 admin
State 2, school 3	Rural	291	7–8	99	23		3	2	
State 3, school 1	Rural	370	7–8	97	17	8% SPED	4	2	
State 3, school 2	Rural	427	K–8	96	8		1	1	
State 3, school 3	Urban district	~550	6–8				3	1	1 science, 1 admin
State 4, school 1	Small town	600	6–8	94	39	3% ELLs	3	2	

Interviews were conducted by four people, based on an open-ended interview protocol. They were then transcribed and analyzed, using qualitative software. Initial codes were developed based on both the literature and extensive meetings held with the project management team. These initial codes included “opportunity to learn,” “learning style,” “test stress,” and “motivation,” among others. Subsequent recoding led to new categories, which were then checked through recoding of interviews. Coding was done by the two senior researchers, who regularly shared coding strategies, coded each other’s interviews, and redid coding based on emerging categories. As a result of this coding process, we developed two major coding categories, “student characteristics” and “assessment characteristics.” These two categories led to the development of two research papers.

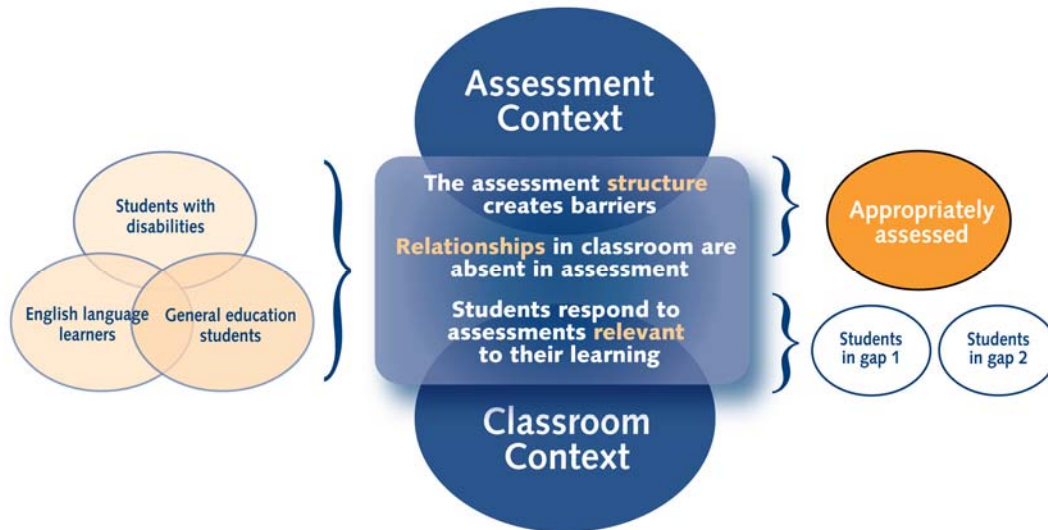
Because we chose to interview only teachers for this study, we recognize that there are limitations in the conclusions we can draw. We specifically sought to understand the teachers’ viewpoint, but by limiting ourselves to this viewpoint, we do not have any data that would indicate teacher bias, such as student interviews, classroom observations, or a systematic analysis of student assessment data.

Findings

Based on the results of the interviews in this study, and together with the other parts of the study, we have developed a theoretical framework demonstrated in Figure 1. On the far left side are three groups of students: students with disabilities, ELLs, and general education students. When these three students are faced with the large-scale assessment, some of them are negatively affected by differences between the classroom and the assessment context. These differences include the structure of the assessment, the differences between teacher practices and the assessment, and the lack of relevance of the assessment to student learning. Depending on the particular attributes of each student, these challenges result in them being affected by either the

first or the second gap in the assessment system, as demonstrated at the far right of the framework.

Figure 1. Explaining the gap between classroom achievement and assessment results



This paper focuses on the students described in the framework. We describe the two gaps that teachers identified, and we look at the experiences of students with disabilities, ELLs, and general education students, from the point of view of their teachers.

The findings in this paper are organized into the following sections. First, we describe the two assessment gaps that the interviewees identified. Second, we look at their descriptions of students with disabilities and ELLs, noting that these students are found in both gaps. We then describe the attributes that teachers identified as common to all students who struggle with assessments, including general education students. Finally, we look at issues of access to the curriculum and out-of-school issues.

Teachers describe two gaps in the assessment system

In our interviews, we asked teachers to describe the kinds of students who tend to perform below proficient on the large-scale assessment. We began with the explicit charge to focus on those students who present a "discrepancy," that is, whose performance is lower than teachers would have expected. We found, however, that in many of the interviews, teachers also wanted to talk about the group of students at the very low end of the spectrum, the ones who were not even learning the material at grade level but who did not qualify for alternate assessments. These students found the large-scale assessment to be a painful experience, and one that simply confirmed what was already known about their lack of proficiency. At the same time, these students are progressing in their learning, although, clearly, they are not learning to grade level standards. Teachers in each of the four states argued that the gap between the alternate and the regular assessment affects not only students who "don't test well," but also those students who

are advancing well but are still below grade level. We identified the gap affecting students who “don’t test well” as gap 1, and the gap affecting those students who consistently obtain the lowest scores on the assessment because they are working below grade level in the classroom as gap 2. Table 2 provides a summary of teacher descriptions of students in the gaps.

Table 2. Teacher descriptions of students in the two gaps

	Gap 1	**Gap 2**
	• Generally taught at grade level • Large-scale assessment does not adequately measure their proficiency, as demonstrated in the classroom	• Generally taught below grade level • Perform at the floor of large-scale assessment because they are being taught, and are learning, far below grade level
Students with disabilities	• Unspecified learning disabilities • High-incidence disabilities	• Cognitive deficits • Severe learning disabilities
ELLs	• Sufficient understanding of academic language to access content areas, but still do not show their proficiency on large-scale assessments	• Minimal understanding of academic language and cannot show their content proficiency on large-scale assessments
General education students	• Different learning styles • Low motivation • Test stress • Out-of-school issues • Attendance issues	• “Slow learners” • Low motivation • Relevance • Out-of-school issues • Attendance issues

One teacher talked about challenges faced by a student who does not fit the criteria for alternate assessment: “There is one child in particular—I’m sure he will never move beyond second grade reading. He is working to his ability. He increases his reading, his vocabulary, within a second grade level, but certainly we can’t put a sixth grade test in front of him. And yet he didn’t qualify, from what our understanding was, for an alternative portfolio.” (NB, SPED) Another teacher summarized:

> They tend to be slow learners. They tend to be “shady 80s,” that is what I call them. Seventy to 75 makes you mentally retarded or learning impaired. If you are in the 90s, then you are okay. These are shady 80s. They show up every day for school, and they sit down and crank out their little homework. They don’t have a clue what the homework means, but they have it done. They always have a notebook. They always have a sharpened pencil . . . These guys are good students in the classroom. They have their notebook and their pencil. Clueless. They have no mechanism to practice it. . . . They never move from that very pretend area of teaching. They do it fairly well. These kids are going to get 70s. These kids will get it right, but they don’t have a clue how they got it right. It never becomes theirs. (CB, SPED administrator)

The first teacher described the frustration of not having an appropriate assessment for her sixth grade student performing at the second grade level. The second teacher focused on the perception of student ability, arguing that the group of students just above the alternate assessment don't learn the material they are presented in class, and have adopted strategies to get through the school day without actually learning. Other teachers focused on the difference between a student's IEP and the large-scale assessment. This contradiction, where an IEP may specify mathematics at the fifth grade level while the assessment measures at the eighth grade level, is a gap that particularly concerns many special education teachers:

> Many of those children are below grade level to begin with. And we are asking them to perform at a sixth grade level. When they are already below grade level, how can they perform at that level? And most of their program during the year is to meet the objectives on their IEP, which addresses the fact that they are below grade level . . . We have several life skills children on our team that are significantly below grade level. And their IEPs would not necessarily . . . have assessed the multiplication to the point that the [state assessment] did. (DT, Mathematics)

The above quote raises a critical question that is not answered in this study but that continues to plague conversations on this theme: Are students being taught to their highest potential? IEPs should be focused on access to the general curriculum. How can it be definitively determined that a student cannot learn at grade level? Indeed, one special education teacher described his changed understanding of student potential:

> My teaching practice just in the past year has changed dramatically. My thrust now is to really concentrate on eighth grade GLEs [grade-level expectations], even though most kids I have in my resource room are third grade level for math, maybe fourth grade . . . I found math strategies presented on the third, fourth grade level . . . they were exposed to strategy on their level, so then we worked through problems up to the eighth grade level. So that's a new direction for me, because I'm not sure I've always had the expectation that they could do eighth grade math. So my expectation has changed, and my teaching practice as a result of that has changed. (BS, SPED)

Thus, while teachers clearly and unequivocally identify a group of students in gap 2, there is no consensus on how to differentiate between the need for different instruction and higher expectations and the need for a more appropriate assessment.

Given this initial finding that teachers identify two distinct gaps, the next step was to look for attributes of the students in each gap. Students with disabilities and ELLs, not surprisingly, often fall in one of the two gaps. However, in the interviews, teachers listed various student attributes that have a negative impact on test performance and that cut across students with disabilities, ELLs, *and* general education students, meaning that the gaps also comprise some general education students. For students in gap 1, interviewees readily explained the correspondence between students' sub-proficient scores and their learning styles, motivation, and self-esteem. Students in gap 2 demonstrate many of the same characteristics, but they present the added challenge of learning below grade level. In short, according to interviewees, the assessment is not an effective measure of content proficiency for either group of students.

Students with disabilities are found in both gaps

Students with disabilities are among the subgroup with the lowest scores on large-scale assessments. In 2004, 66% of New Hampshire students with disabilities scored in the "below basic" category in sixth grade mathematics, and 75% of Maine students with disabilities scored in the "does not meet standards" category in eighth grade mathematics. The teachers and administrators interviewed for this study all agreed that students with disabilities continue to struggle with demonstrating their proficiency on large-scale assessments.

While many students with disabilities fall into our gap 2, other students with disabilities are working closer to grade level and fall into gap 1. They are being taught at grade level and perform at grade level in the classroom, but do not demonstrate that proficiency on the assessment. Teachers described factors that limit the ability of these students to demonstrate their proficiency:

> Even with the modifications that the state allows us to do, oftentimes putting pen to paper or pencil to paper—or just the whole magnitude of taking of tests of that magnitude, [which] puts kids at a disadvantage who have some type of handicapping condition. And it also truly puts them at a disadvantage, because they don't have the self-confidence, either. They're being taken out of the classroom, away from their peers. So automatically, everyone knows that they're not of the same caliber as the rest of the kids. (LM, Mathematics administrator)

Short-term memory

Teachers identified specific challenges that students with disabilities face in eighth grade mathematics, including short-term memory. One teacher commented:

> It is very difficult, I think, to expect kids, particularly at that middle and lower level, to remember everything from September that we've done, even though my program reinforces all the way back up to March. Maybe a portion could be taught and tested and then move on to another portion. The test covers everything. And it just seems way too much for those kids. (BB, Mathematics)

This idea that a single annual assessment is not an effective tool to measure the learning of all students emerged in many interviews. Teachers consistently talk about students struggling with recall. Sometimes students know that they have seen a concept before, but they need to have the teacher give a few examples to remind them of the process. Students need to "use it all the time," but that is impossible to do with all the math standards. One teacher noted:

> It seems like you move on down the road, and two months later you mention something you have done, and you get these blank looks and like, "I know we covered this. I can look in the plan book. Remember when we did this? I said this and this." And then the little lights come on, "Oh, yeah, you did do that." (BW, Mathematics)

Issues of short-term memory are particularly critical when looking at large-scale assessments, because large-scale assessments typically measure a full year's worth of learning. If students learn and understand a concept in September, but when asked to apply that concept in March

they can no longer remember it, what should be done? Have they actually learned the concept if they cannot recall it six months later?

Focus

Students with disabilities tend to struggle with the need to focus for long periods of time. One teacher noted, “This is a girl who scored 200 because she literally was so off-task that even with constant reminders, ‘Please, try to go on to the next one,’ she was just—she would spend an hour and answer one question. She was so off-task.” (PT, SPED) For some students, it is hard to focus on the multi-task orientation of an assessment. In particular, teachers described students who have a hard time maintaining focus for long periods of time. A teacher from Vermont related, “If they burn out, they burn out with assessments being that long. We have one and a half hour-sessions for the eighth graders, with eight sessions. You are going to lose them early on.” (FL, Mathematics) Another echoed this concern: “That long period of time, concentrating, and sitting there, when they have an attention span of two and a half minutes anyway. Classes are usually 45 minutes long. We usually have three or four shifts of what we are doing in that time because they don’t attend all that well. When you give a test, and it is one thing for twice that period of time, it is a pretty different scenario.” (CW, Mathematics)

Thus, for students with disabilities, accommodations do not always address their needs, and short-term memory and focus issues interfere with their ability to demonstrate proficiency.

ELLs are found in both gaps

ELLs make up more than 10% of the student body in only two of the schools in our study. Thus, in this study, the challenges facing ELLs did not emerge as strongly as they would in areas with larger numbers of these students. However, at the schools where there are ELLs, or where a teacher specifically mentioned ELLs as struggling with the large-scale assessment, we sought more information about teachers’ perceptions of the students. The teachers identified ELLs as fitting into both gaps. Students who have been studying English for only a short time, for whom “the language of the test has no meaning,” fall into the first gap. Yet, unlike other students in gap 1, they may know the content; they are simply unable to access the test. Students who have learned more English are more likely to be in the second gap. One math teacher noted, “I have another student that is an English language learner and is new to our school. And he failed the test and came to me. And we sat down then for 10 minutes, just one on one, and [I] said, ‘Do you understand why this needs to happen in this way?’ He totally got it.” (JM, Mathematics)

A math administrator from a district with a growing ELL population, who had worked in one school for many years and just recently became a district administrator, noted how difficult it is to prepare ELLs for the large-scale assessment because of the inability to pre-teach: “Normally what happened with the ELL teacher in this school would be, a math teacher might give the vocab[ulary] words ahead of time, so that the child can at least have an even plain when the lesson starts. With the [state assessment], you can’t do that.” (LM, Mathematics administrator)

The same administrator recalled an experience with one of her students: “I have seen kids in the ELL program—one in particular, who completely broke down. And this is a child who had pretty much normal capabilities. But because of post-traumatic stress, living in a different country, language limitations, new school, really, the [state assessment] put her over the edge. And the

ELL teacher was part-time, and so the relationship wasn't able to foster what needed to be fostered."

Other student attributes contribute to assessment gaps

When we began our research, we brought to it an assumption that teachers would generally identify students with disabilities or ELLs as the students who struggle the most to demonstrate proficiency on large-scale assessments. However, preliminary analysis of the large-scale assessment results in two of the states in the study (Maine and New Hampshire) show that the universe of students who do not reach proficiency goes far beyond these two subgroups.[1] In New Hampshire in 2004, 66% of sixth grade students with disabilities scored below basic on the large-scale assessment, a total of 1,718 students. Among students without disabilities, 20% scored below basic, a total of 2,880 students. Thus, while on average general education students were more likely to score in the proficient category, the absolute number of general education students scoring below basic is higher than the absolute number of students with disabilities.

Trends were similar in Maine for eighth grade students taking the state assessment in 2004. While 75% of students with disabilities scored in the "does not meet the standards" category (1,939 students), 24% of general education students also scored in that category, a total of 3,476 students.[2] Both New Hampshire and Maine have large numbers of general education students who are performing below proficient on the large-scale assessments. Thus, the numbers show something that we found in the teacher interviews: When teachers look around their classrooms at the students who are not meeting proficiency on the large-scale assessments, they see students with disabilities, ELLs, *and* general education students.

The teachers in our interviews, when asked to describe the classroom behavior of students who struggle with the large-scale assessments, described a wide range of characteristics and behaviors. Students with learning disabilities tend to experience challenges with focus, with organization, and with multi-step problems, as described above, but teachers also describe students *without* disabilities who have all of those characteristics. Students who struggle with motivation are found among the students with disabilities and *also* among the students without disabilities. The following section describes some of the issues raised by teachers, and illustrates how the teachers find these issues in all sorts of students, far beyond the standard categories of students with disabilities and ELLs.

Why might there be students without identified disabilities who do not demonstrate proficiency on the large-scale assessment? One big issue may be instruction—either the students have not had the opportunity to learn the material because teachers have not presented it, or the teachers may have taught it but the students need more individualized instruction even if they don't have an identified disability. Another issue may be the process by which students are identified with disabilities. A mathematics teacher noted, "I have a student who is low IQ, low performing. So she is non-special ed. And she needs to be taken out, and she needs to either do it piecemeal or have extra help." (BW, Mathematics) One special education teacher explained that a number of students who are not identified with a disability have learning needs that need to be addressed:

---

[1] Data were not available for analysis in Rhode Island and Vermont.
[2] The numbers of ELLs in the two states is far smaller, though patterns were similar.

"When you have borderline MR [mentally retarded] kids, then you are in trouble. I mean, you know, if their IQ is in the 70s and they are not really MR, then they don't get identified. Or even 80s. It is hard for those kids." (PR, SPED) In another school, a pullout math class was developed for students lacking basic math skills. Half the students in the class had identified learning disabilities, but the other half did not. The teacher explained:

> I also teach a math class for students who don't do well on the testing, for kids who don't have the basic math skills, math facts down. They are capable of learning but not at the fast pace of the classroom . . . As a matter of fact, I had nine children this year. Four were identified as learning disabled and five were identified as . . . more in a slow learner range. Like one of my girls, she basically learned almost everything the other kids learned in mainstream (because they are all self-paced when I take them) but just could not, the spiral focus that they use, before she could grasp it, they moved on. And it just didn't match her learning style. (NB, SPED)

As we reviewed the teacher interviews, we found that while the descriptions varied widely, we were able to identify three major areas where student attributes had an impact on assessment results: learning styles, motivation/interest, and content knowledge. The following section describes the ways that teachers explained the role of each of these three areas in student demonstration of proficiency on large-scale assessments.

Different learning styles

Students learn in different ways. Some students prefer to see everything visually, others are more hands-on, while others are auditory. One teacher explained, "The tests, the way they are constructed, are very problem solving, very verbal. More verbal kids, the ones that don't have any of the verbal issues, tend to not fear them and tend to do all right with them . . . So the kids that are more visual, who are more hands-on, who need to know, struggle more with them than the kids who are very verbal and have good decoding skills." (CW, Mathematics) The teachers interviewed identified the way that some learning styles are more useful for a sit-down assessment, while other styles lead to frustration, such that students do not demonstrate their proficiency. They also noted some areas not often included in the literature about learning styles, including abstract thinking, independent learning, and test stress.

Abstract thinking

Teachers noted that some students are stronger abstract thinkers than others. One teacher argued that the assessment demands more abstract thinking than is typical for middle-schoolers: "Abstract thinking is something that comes developmentally. It doesn't come by the seventh grade. Developmentally, you are lucky to reach that in your mid to late teens." (SL, Mathematics) Another described a typical classroom experience of having students not understand directions on a classroom test, even though they have used the exact same directions for homework assignments: "These kids expect that whatever we practice is exactly what they are going to see later. They don't make the relation of a different problem to the one they did unless they see almost exactly the same numbers." (BB, Mathematics)

Teachers who describe students in developmental terms argue that not all eighth grade students are ready to be assessed on eighth grade material at the same time. Their argument is not that

these students will *never* reach grade level, but rather that not all students follow the same strict timeline:

> The reason why I like you asking these questions is that the middle school kid is a really hard kid to know. When they grow from concrete reasoning through that whole abstraction thing, it is really hard to know when you are assessing them where they stand in that spectrum. If you use a test to test them, are you getting an answer that has something to do with their test-taking ability or with their abstraction abilities? It is really hard, and we struggle mightily in this school with how to place kids in different classes to meet their abilities. (CW, Mathematics)

Teachers talk about students who are more interested in the changing social dynamics of middle school, or more interested in having fun. They tend to believe that these students are developmentally not ready for the abstract work of algebra, but that over time the students will probably catch up.

Independent learning

Independent learning is similar to abstract thinking, in that teachers describe it as being an important element in success in mathematics. According to one teacher, "Some of the folks that we have that you define as [students in the gap], I would view them as kids that do not take risks on any number of levels. Like, anything that looks new or unfamiliar, they are not going to go back in their mind to think about prior knowledge." (MG, SPED) Later in the interview, when talking about a specific experience with a group of students, she added, "They just could not deal with the ambiguity. I find that the students that do better are the students that ask questions. The students who do better are the ones that are willing to make associations between things." (MG, SPED)

Those students who are independent learners, who are willing to take risks, are more likely to solve a problem that doesn't look exactly like one they have solved before. They are students who are more likely to recognize a problem in an assessment setting that they might have seen before in a classroom setting. This characteristic is related to the third learning style that teachers described: a tendency to experience stress around test-taking.

Test phobia/stress

Not surprisingly, a number of teachers mentioned stress as being an important factor that keeps students from demonstrating what they know on a large-scale assessment. One teacher summarized it well: "You are always going to lose kids because of one variety of problem or another. Either they have test anxiety and freeze up so they can't focus, or you have kids who cannot get what they have inside onto the paper. You will always lose those kids in a paper-and-pencil environment." (RV, Mathematics) A teacher from Vermont noted, "A lot of students, even students that are not special ed, have a bad case of test phobia—so they do really well in the classroom, but on that day when they have to take any kind of a test, standardized or any assessment at all, they just freak out and can't do it." (GG, SPED)

Similarly, another teacher described the difference between a student taking a test and the same student explaining what he knows:

> But I [can] walk through the book and have him show me. I'd say, "How would you do this," and he would show me. But if I gave him a test on it, I know—he does terrible on tests, terrible. I mean it is like—and if I take the test with him, and I say, "Okay, let's take a minute and look at this particular one." And he has an answer. "How do you answer that?" And he doesn't know how he answered it. (KD, Mathematics)

Not all teachers agree that students are affected by stress. One teacher noted, "I see it more as a great excuse. I don't know. I don't see any kids really stressing out terribly during a test." (SL, Mathematics)

Motivation and/or interest

Teachers made direct links between motivation and previous learning. Six teachers specifically referred to the negative impact of past failures on current achievement and motivation. This teacher described the negative cycle well:

> You know, they are so far behind that they just don't see a light at the end of the tunnel—"Why should I try?" . . . They have been failing so often. And instead of pulling them aside and being able to . . . regroup and really focus on what it is they need, so that they can move on to the higher skills, [there is] not enough money, not enough people, in some cases not enough motivation on their [part]. (JT, Mathematics)

Others described the way that early failure has a negative impact later:

> Not so much testing history that I have seen, but more history with the printed word. Failure with the printed word, even at an early age. A lot of kids who had early failure, even though they have moved on and are really capable of doing it, they hold on to that original idea as who they are as learners. It is sort of like the old saying: "There is a man who thinks he can and a man who thinks he can't, and they are both right." I think a lot of healthy kids become the man who thinks he can't. (MC, SPED)

Motivation is critical because it provides the impetus to do the hard work of learning. Most students have to contribute to their learning by putting in the effort. Teachers describe students for whom the primary limitation to success is their own effort. One teacher had some students in this category: "And I felt in a way then, the assessment was an accurate measure of what they could do because they never did anything. So when they wouldn't engage in class or any of the classroom material, or any of the work all along, they ended up doing very poorly on the test because they didn't do anything, all along." (JM, Mathematics)

Other teachers are less generous in their assessments of their students: "Kids don't pay attention. Kids just plain don't pay attention. I can be going over what we are doing, and I can say [that] literally a solid 20% of the kids don't hear the words that come out of my mouth, all day." (KD, Mathematics) However, most teachers believed that students experience a combination of motivation and previous knowledge that combine to make test-taking difficult: "The ones I saw give up were the borderline ones that may have been able to pass it if they absolutely gave it their 100%." (RH, Mathematics) Another teacher echoed this: "And we are also talking about the

group of students that might be the average student to the low-ability student, so they aren't quite as willing to work so hard, aren't as willing to practice, practice, practice until they do get it. The competitiveness is not there." (BW, Mathematics)

A number of students end up missing large chunks of schooling, whether because of truancy, illness, or behavior issues. For one teacher, a number of her students performing at the lowest level were chronically absent: "There were definitely some kids that were way at the bottom. And one of the big issues there is absenteeism. That would definitely be one of the reasons for poor performance—continued absenteeism." (JM, Mathematics) Another teacher highlighted a student with behavior issues who missed class when he was kicked out frequently:

> He is a boy who I think is really, really, really smart. And he was able to do a lot of computation in his head. But he was an extreme behavior problem. He was often kicked out of other classes. . . . But I was really curious as to how he would do, because typically on report cards, he wouldn't do well, because he would never do homework. He would oftentimes act up in my class, and his class work was a part of his behavior. So it didn't really give a good look as to what his performance was, even though I thought he was a really sharp kid. So in a nutshell that was one particular kid. (BN, Mathematics)

Teachers identified self-esteem as being an important element in student learning, arguing that low self-esteem has a negative effect on achievement. One math teacher noted that there also seems to be a "genetic thing" involved:

> In math, especially in kids—reluctant learners—they come with all kinds of baggage, from not feeling confident to truly not having mastered the basics to continue with a math education. In some cases, they think that their not being able to do math is a "genetic pool" thing. Their parents weren't good at it, so therefore they're not good at it. (LM, Mathematics administrator)

Teachers also found that they could develop ways to address self-esteem issues with their students. One teacher talked about the benefits of providing moral support:

> As long as they are in the classroom and they have an aide or a teacher that shows them how to do it, then they feel like they are understanding it, and they do really well. I have one student that as long as you are sitting next to her, she can fly through her work. If you get up and walk away from her, she doesn't have a clue what to do. You can just sit there. You can read the paper and not say a word. As long as you are sitting there, she works well. (GG, SPED)

Vignette: An alternative learning setting

At one school in Maine, the principal designed an alternative setting specifically to address the needs of students who are at risk of dropping out of school—not because of a learning disability or behavior problems, but because they have not learned at grade level, and thus are rapidly disengaging from learning and are more likely to drop out. They are students who, as described above, struggle because they did not learn material in earlier grades, they do not put in any effort, and they have high absentee rates and low self-esteem. For one year, nine eighth-graders were taught all subjects by one teacher in a single classroom. This teacher found that by targeting instruction to each student, almost all of them reached grade level in all subjects by the end of the year. We interviewed the teacher at the end of the year, and she reflected on where the students were when the year began, where they are now, and how they got there:

> Most of them, in math, were at the fourth grade level, and they were eighth-graders. Not because they could not do the work. They sat quietly in a classroom and just vegetated . . . They didn't necessarily act out, some of these kids . . . They know how to behave so a teacher won't call on them. They know how not to pass in a paper because they don't [want] to discuss with anybody why they don't know how to add or subtract fractions. It is better for them to say "I forgot my homework" than to deal with the real issue. Their self-esteem was zero. So, anyhow, these are the things that we discovered. (BD, Alternative education)

The teacher emphasized that the students in her alternative class did not have learning disabilities: "There is nothing wrong with their ability to learn. Nothing. I mean they are not—they don't have any learning disabilities. They just are lazy and haven't done it. Once they started, they started to move." (BD, Alternative education) She spent the year working with each student individually, breaking through the issues that had kept them from learning:

> What we found out was [that] they have these misconceptions, but they never articulated them. They didn't want to appear stupid. So they have their own little ways of doing problems. And when I will say to them, "I'd like to show you how to do this"—"No. No. No. I'm showing you." And I said, "Well, here's the deal. Your way isn't successful. You have an F. So I think maybe you need to get rid of your way and look at my way and see if it works," that sort of thing. So we have this one on one. And they would rather do all the problems, hand them in, and take the F, too. That's another thing. And I won't let them. So what has worked for me is, "You do the first half of the paper. You turn it in. I spot-check, see how you are doing. And if you have a problem, I'll show you what to do. You go back and redo it." (BD, Alternative education)

The teacher was enthusiastic in her evaluation of the students' progress. In each subject area, they had advanced far more than one year's worth, and in one situation, the home relationship to school also changed: "This [girl], whose parents never spend time with her, her mom stayed up and drilled her, asked her questions from flash cards. So that was, even the interaction between mom and daughter was good." (BD, Alternative education) The school targeted instruction to meet the specific needs of these students, though there was no opportunity to do the same with the state assessment.

Access to grade-level curriculum

As noted above, we began the research with a certain hypothesis about the gaps in the assessment system. We hypothesized that teachers would describe students who demonstrate proficiency in the classroom but not on the assessment. Indeed, teachers did talk about those students, but they also focused far more on another group of students: those students who are above the alternate assessment, but who are being taught, and are performing, below grade level. Every teacher we interviewed talked about these students. Some of these students have identified learning disabilities; others do not. Some are in the regular classroom; others are in special education classrooms. This section highlights two areas of particular concern among teachers for these students in gap 2: mastery of basic mathematical concepts and perceived ability to learn at grade level.

While all teachers described students struggling with basic skills, mathematics teachers in particular identified the gaps in student learning. Some teachers describe a lack of skills, arguing that students need to be taught these skills in different ways, while others argue that some students will never grasp the mathematical concepts needed for pre-algebra and algebra. One mathematics teacher explained:

> In terms of math, last year I was very surprised at the lack of basic skills that my eighth grade students came to me with. And maybe this is wrong, but I felt that they had not been held to a standard that they had to show that they had mastered skills. So that when they were given these assessments, though they had the concept of multiplication, they had never learned their multiplication tables. (JM, Mathematics)

The teachers inadvertently addressed a critical issue: teacher expectations for students. A recurring theme throughout this research has been the debate about student ability. Are some students simply not capable of learning to grade level? Or, if teachers held higher expectations, and if instruction were more targeted to individual needs, would students learn more and perform better? As teachers talked about mathematics content, it became clear that some of them find that students need different, or more, instruction. Again and again, teachers say both that students need more instruction and that the assessment is not a good measure of what some students know. The students find themselves in a gray area between the need for a different assessment and the need for different instruction, or perhaps both.

According to the teachers, students in gap 2 were not only students with disabilities or ELLs. A number of teachers described working with general education students whose mathematics skills were far below grade level. Interestingly, in a number of schools, special education teachers described working with the general education students because their schools targeted student support based on needs rather than on a specific designation:

> But the basic skills still have to be there, and these kids have no skills. I mean, it is scary. I mean I, literally, talking about a kid that didn't qualify for special ed—she was in my social studies class back a few years ago. She could not read. Could not do single-digit addition. And she is in seventh grade. How can she perform in society if she can't do that

> by the seventh grade? I mean, she is a sweet kid, a nice girl. I went as far as to have a parent volunteer do the textbook on tape, so at least she could hear it. I mean, there is no way she would pass a regular class. (AT, SPED)

Teachers describe students who can grasp mathematical concepts but who struggle in mathematics because they lack computational skills. One teacher said that students just "throw math at a thing" (KD, Mathematics), without knowing how to figure out which operation should be used. While most teachers talk about strategies they use, or would like to use, to improve students' understanding of math sense, others have found that they have some students who just "don't get it." These teachers believe that some students are not math-minded and won't learn:

> No matter how many times or how many different ways I present it, they don't understand it because they don't understand how numbers work together and how they fit. They see no patterns. To them, it is just a bunch of numbers. They have never made that connection before. So I end up trying to "teach to the test" for that number of students. (BB, Mathematics)

When talking about students who they believe cannot learn at grade level, teachers describe both specific mathematical challenges and general learning challenges. Earlier in this paper we described the challenges teachers face when there is a contradiction between specific student IEPs and test-taking requirements. In addition, other teachers described students without disabilities who don't have IEPs, and so should be working at grade level, but who are being taught, and are learning, below grade level:

> And another student that I'm looking here that did not meet was low-functioning in all her classes, but was never diagnosed as having anything. But it really was very obvious to us that either she had a very, very low IQ or [she] indeed needed some . . . [support] that we didn't have the resources to provide. (JM, Mathematics)

Thus, gap 2 students can have all the characteristics that describe gap 1 students, such as different learning styles and motivation issues, but these differences are compounded by being far below grade level. These students take the same assessment as all other students, even though it is clear to teachers that they cannot reach proficiency, either because they are being taught below grade level in the classroom, or because they have not demonstrated mastery of any grade-level concepts or skills.

Out-of-school issues can affect assessment results for all learners

For teachers in the classroom, the daily realities that students face in their out-of-school lives play an important part in student performance. In some cases, teachers described the ups and downs of being a middle school student. In others, they identified particular crises that affected student performance. In still other cases, teachers suggested that student poverty and/or parental engagement have a strong effect on assessment outcomes. Education and assessment policies cannot change the out-of-school experiences of students, but they can change the way that teachers, schools, and the assessment system take account of them.

Some teachers talk about the importance of day-to-day student experiences and the impact on large-scale assessments: “Eighth-graders are a pretty interesting breed. On any given day, a brother and sister could have gotten into a fight that day and, ‘My day is upside down. I can’t be focused on these numbers today.’ It is such a clouded reality.” (RH, Mathematics) One student had to deal with world events: “This child, her dad was in Iraq. She ended up just about failing the class. That kind of thing comes up.” (SL, Mathematics)

Teachers described poverty as a challenge that many children face, and often linked poverty to engagement with the educational process. Others criticized [lower income] students and their families for not having an attitude conducive to learning. One teacher commented, “Students who are from lower [socioeconomic] groups tend to not be as engaged. They seem to have [an] attitude that has been instilled in them that education is not valued and is not important. So they tend to have a negative attitude. If your parents have a negative attitude toward education, then many times [the students] do as well.” (FL, Mathematics) Another commented, “The kids that don’t really value the process are from the lower end of the spectrum.” (MC, SPED) One teacher put it a little differently: “They just don’t get any cheerleading in their life, and they come to it thinking they are going to fail.” (CW, Mathematics)

The teachers in the interviews, while acknowledging the challenges presented by out-of-school issues, and in some cases almost blaming the students and their families for low performance, also described ways to address these challenges. Many of them talked about the importance of relationships in mitigating some of the complicating factors that students bring to school with them. A mathematics teacher, noted how large-scale assessments don’t “assess the fact that these kids are making relationships.” According to this teacher, the test “misses” key components of student experiences, those very components that help to address the out-of-school challenges that students face.

Conclusion

This study of teacher views on why students do not reach proficiency in eighth grade mathematics revealed important aspects of both the students and the assessment, and has raised important areas for further research.

The first finding is that teachers identify two different groups of students in two different gaps. In gap 1, students demonstrate proficiency in class but not on the assessment. Teachers describe students with disabilities, ELLs, and general education students in this gap. They argue that the difference between the classroom and the assessment contexts is detrimental to these students. In gap 2, students perform below grade level but demonstrate progress that is not measured in the assessment. Again, all three types of students (students with disabilities, ELLs, and general education students) are found in this gap. Students in gap 2, in addition to being adversely affected by the structure of the assessment, can also be adversely affected by the grade level being assessed.

The second finding addresses the types of students found in the gaps. Teachers noted that in addition to students with disabilities and ELLs, students with no identified disability or English language learning issues fall into the gaps. When teachers look at a classroom of students, those

who are poorly assessed may or may not have an identified disability or language need. More research should be done looking at these students and their instruction and assessment needs.

In both gaps 1 and 2, this study could not definitively delineate when the gap is truly an assessment gap that should be addressed through assessment changes, and when the gap is instruction-based. Teachers said that in many cases, the mathematics assessment does assess students' levels effectively, and those students who do not reach proficiency do not, in fact, know the material. Thus, the assessment is doing what it is supposed to do. This leaves us with the unanswered question of how much of the gap lies with the assessment, and should be addressed by changes in the assessment, and how much lies with instruction. The teachers in this study, while arguing that the large-scale assessment fails to account for many facets of student progress, also tended to agree that the mathematics assessment is a good measure of the standards it claims to assess. This study was only able to subjectively identify those students who would benefit from a different assessment rather than from different, or more, instruction. As noted in our vignette about the alternative school, some students, when provided with a learning environment tailored to their needs, can make amazing gains in a relatively short time.

Future research studies focused on students in the gaps should seriously consider conducting in-depth interviews with the students themselves, to better understand the ways that students understand the assessment and the gaps that affect them. In particular, this research was unable to sufficiently explore the issue of relevance for students. How do students understand the relevance of large-scale assessments that are used to determine adequate yearly progress at the school level? What are the issues of relevance related to content, and what are the issues of relevance related to the test-taking process itself?

References

Abedi, J., Courtney, M., & Leon, S. (2003). *Effectiveness and Validity of Accommodations for English Language Learners in Large-Scale Assessments.* Los Angeles: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards and Student Testing.

Abedi, J., & Dietel, R. (2004). *Challenges in the No Child Left Behind Act for English Language Learners.* No. Policy Brief 7. Los Angeles, CA: CRESST.

Almond, P., Quenemoen, R., Olsen, K., & Thurlow, M. (2000). *Gray Areas of Assessment Systems.* NCEO Synthesis Report No. 32. Minneapolis: University of Minnesota, National Center on Educational Outcomes.

August, D. (2006). *Developing Literacy In Second-Language Learners: Report of the National Literacy Panel on Language-Minority Children and Youth* Washington, DC: National Literacy Panel on Language-Minority Children and Youth.

Bielinski, J., Thurlow, M., & Ysseldyke, J. (2001). *Read-Aloud Accommodation: Effects on Multiple-Choice Reading and Math Items.* Minneapolis: National Center on Educational Outcomes.

Bolt, S. E., & Thurlow, M. (2004). Five of the Most Frequently Allowed Testing Accommodations in State Policy. *Remedial and Special Education, 25*(3), 141-152.

Dolan, R. P., Murray, E. A., & Strangman, N. (2006). Mathematics Instruction and Assessment for Middle School Students in the Margins: Students with Learning Disabilities, Students with Mild Mental Retardation, and Students who are English Language Learners. Draft Paper preapred for the Rhode Island Enhanced Assessment Grant.

Edgemon, E. A., Jablonski, B. R., & Lloyd, J. W. (2006). Large-Scale Assessments: A Teacher's Guide to Making Decisions about Accommodations. *Teaching Exceptional Children, 38*(3), 6-11.

Exceptionality. (2004). Large-Scale Testing of Students with Disabilities. *Exceptionality, 12*(2), 67-70.

Francis, D. J., Rivera, M., Lesaux, N., & Rivera, H. (2006). *Research-Based Recommendations for the Use of Accommodations in Large-Scale Assessments*. Houston: Center on Instruction.

Helwig, R., & Tindal, G. (2003). An Experimental Analysis of Accommodation Decisions on Large-Scale Mathematics Tests. *Exceptional Children, 69*(2), 211-225.

Hofstetter, C. H. (2003). Contextual and Mathematics Accommodation Test Effects for English-Language Learners. *Applied Measurement in Education, 16*(2), 159-188.

McCardle, P., Mele-McCarthy, J., & Leos, K. (2005). English Language Learners and Learning Disabilities: Research Agenda and Implications for Practice. *Learning Disabilities Research and Practice, 20*(1), 68-78.

Quenemoen, R., Thompson, S., Thurlow, M., & Olsen, K. (1999). *Forum on Alternate Assessment and "Gray Area" Assessment.* Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Shaftel, J., Yang, X., Glasnapp, D., & Poggio, J. (2005). Improving Assessment Validity for Students with Disabilities in Large-Scale Assessment Programs. *Educational Assessment, 10*(4), 357-375.

Thompson, S., Johnstone, C. J., Anderson, M. E., & Miller, N. A. (2005). *Considerations for the Development and Review of Universally Designed Assessments.* NCEO Technical Report No. 42. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Universally Designed Assessments. (2005). Retrieved May 5, 2006, from http://education.umn.edu/nceo/TopicAreas/UnivDesign/UnivDesign_topic.htm

Wiener, D. (2006). *Alternate Assessments Measured Against Grade-level Achievement Standards: The MA "Competency Portfolio".* NCEO Synthesis Report No. 59. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Ysseldyke, J., Thurlow, M., Bielinski, J., House, A., Moody, M., & Haigh, J. (2001). The Relationship Between Instructional and Assessment Accommodations in an Inclusive State Accountability System. *Journal of Learning Disabilities, 34*(3), 212-220.

Appendix A. Interview protocol for initial teacher interviews: Looking at assessment gaps

Think of the students in your school who did not reach proficiency on the eighth grade math state assessment.

1. Do these students fall into any particular categories? If so, what are they?

(If students who don't reach proficiency but who demonstrate proficiency in class are mentioned, direct participants to this group. If these students aren't mentioned, mention them.)[*]

Think of students in your school whose class performance indicates that they really do have the knowledge and skills to reach proficiency on the state test, but whose assessment results are below proficiency.

2. Can you describe these students for me?

    Gender

    Age

    Ethnic background

    SES

    Second language issues

    Learning issues/SPED

    Learning style, strengths, weaknesses

    Academic history—classroom settings, curriculum they were exposed to

    Opportunities to learn

    Family history/background

    Test anxiety/previous testing history

    Attitude toward math, math phobia

[*] We initially targeted only students in gap 1, but teachers consistently described students who fit into a different assessment gap, which became our gap 2. Teachers answered the questions for both gap 1 and gap 2 students.

Attitude toward giving up early

Classroom settings—any special help provided for math or for test preparation

3. On what classroom evidence do you base your belief that this student should have demonstrated proficiency on the state assessment?

   How did you collect this evidence?

   To what state performance level (proficient, beyond proficient, etc.) does this student's classroom performance correspond?

   What kinds of grades does this student get in math classes? How did this student perform in math in school? What are their strengths in math?

   To what kinds of instruction or activities do this student respond best?

   To what kinds of assessment types do this student respond best?

   In your opinion, what about the state assessment makes it difficult for this student to demonstrate what they know and can do?

   What is it about classroom instruction that enables this student to show what they know and can do?

   What would be a better way of measuring what this student know and can do?