

**FINDING A PATH
IN A HIGH STAKE'S
ENVIRONMENT:**

One City Elementary School's Experience with State
Standards and Testing



Reforming Teaching/Learning in a High
Stakes Testing Environment

Capital Region Science Education Partnership
University at Albany, SUNY



Reforming Teaching/Learning in a High Stakes Testing Environment

Capital Region Science Education Partnership
University at Albany, SUNY

Sandra Mathison, Ph.D.
Project Director & Professor
University of Louisville
smathison@louisville.edu
(502) 852 0616

Melissa Freeman, Ph.D.
Project Manager
University at Albany, SUNY
crseptst@csc.albany.edu
(518) 591 8544

© 2003 by Sandra Mathison & Melissa Freeman

Freeman, Melissa

FINDING A PATH IN A HIGH STAKE'S ENVIRONMENT:
One City Elementary School's Experience with State Standards and
Testing
Albany NY: Capital Region Science Education Partnership,
University at Albany, SUNY

This publication is based on research supported by the National Science Foundation (Grant # ESI-9911868). The findings and opinions expressed herein do not necessarily reflect the position or priorities of the sponsoring agency, the Capital Region Science Education Partnership or the University at Albany, SUNY.

**FINDING A PATH
IN A HIGH STAKE'S
ENVIRONMENT:**

**One City Elementary School's Experience
with State Standards and Testing**

Table of Contents

The School	1
Elementary Level State Tests	5
The Educational Climate	20
Parental involvement	22
Instructional time	23
Competition	25
The Trickle Down Effect of Testing	26
Altering practice	
In search of a blueprint	27
Professional development	30
Knowing the tests	32
The push that we needed	33
Teaching the curriculum	33
Teaching to the test	36
The pre-testing years	42
Supporting change	
All eyes on ... fourth grade	44
Meeting the needs of students with special needs	46
Early literacy assessments	53
Going the pendulum	55
Unraveling the Logic of Testing	57

Tables and Figures

Table 1: Willow Valley Elementary scores on NYS 4 th grade tests.....	6
Table 2: Number of new referrals to special education	51
Table 3: History of district and state assessments given to Willow Valley students by grade level	53
Figure 1: Series adopted by Willow Valley School District	28
Figure 2: 2 nd grade story web for Bully Trouble	44

Finding a Path in a High Stake's Environment:
One City Elementary School's Experience with State Standards and Testing

Good Morning!
Today is May 23rd, 2002
Today in history Benjamin Franklin invented the bifocal
Hand in Kids of Character by Friday
The copy room is closed this morning
The food drive continues
Library continues to be held in classrooms
Today's Wizard questions are:
For 4 – 6: Which baseball position can receive a Cy Young award?
For K – 3: What holiday falls on March 17?
Please stand for the pledge
Have a great day!

From the principal's announcements May 23rd could be any day in the lives of administrators, teachers and students at Willow Valley Elementary. But for the school year 2001-02, May 23rd is the first day of the performance section of the science test and the final state mandated standardized test fourth graders have to take. They have already completed three sections of English Language Arts, three sections of math, and the objective section of the science exam. Unlike the ELA and the math tests this one does not receive a space in the morning announcements. Why is that? What are the state tests intended to show? And what do teachers and administrators believe they should do to meet the expectations set forth by the state? This is the story of one elementary school's response to New York State's standards and testing requirements. It weaves together a complex tale of a system's capacity to take on change and the intended and unintended effects of those efforts.

The School

As you enter Willow Valley Elementary¹, the first thing you notice is the exceptional quality of the student artwork displayed around the lobby walls. Vibrant sunflowers; pink, black, and orange Matisse-like collages; political cartoons; or impressionist paintings welcome you as you make your way past the sign-in desk and into the body of the school. As you walk through its multiple corridors, you are likely to encounter rows of children coming to and from their lunch or special, some talking too loudly, some waiting for others or holding hands, others pushing or poking disagreeably. In a corner of the lobby, on the way to art, a little boy jumps down a short stairway and is told to "go back up and do it again walking." He does, only to leap a little less obviously. Down an upstairs hallway, two fifth graders are 'testing' kindergartners' ability to tie their shoes. Two by two the kindergarten students come out of their classroom to demonstrate this ability. If they succeed the older students record their names and place a

¹ Willow Valley is a pseudonym for the school and district

check next to them. If they don't, the older students model the process until they think the kindergartners can do it alone. With over 800 students, four lunch periods in a row, and daily specials like gym or music, moving children from one space to another is part of the daily routine.

Willow Valley district has an enrollment of about 1500 students and includes one elementary school, grades K through 6, and one junior/senior high school. The elementary school services around 825 students. Demographically it is 85% White, 11% African American, 3% Hispanic, and 1% Asian. A free and reduced lunch rate of 61% is evidence of the kinds of needs this school addresses.

Twenty-five years ago a decision was made to close three neighborhood schools and become the only elementary school in the city. Formerly the executive offices for a large abrasives company, the school district was able to purchase the two-story brick building complex when the company moved to another location. Later on a second gym was added as well as a new wing causing the school to spread out in several directions from the large lobby that greets people. With 36 classroom teachers, 8 special education teachers, 5 special area teachers, 6 remedial help teachers, 19 classroom aides and assistants and about a dozen other staff members, the organizational structure of this school is quite complex.

Then ten years ago, as the school population kept growing, the current principal organized a committee to look at the structure and consider reorganizing the K – 6 classes into a schools-within-a-school model. A teacher describes this decision: *“I think the principal was looking for a way to take a big building which was getting bigger by the minute and make it more personal, make it more manageable, 860+ kids is pretty hard to keep track of.”*

The mini-school committee met on a regular basis for 18 months to study and plan how to implement the new model. The intent of the mini-school concept is to promote ownership and to provide smaller groups of teachers an opportunity to work collaboratively on the development of consistent instructional, organizational, behavioral and evaluative plans within their respective mini-school. *“That process was truly one of the first shared decision making processes before shared decision making was a word that we used,”* a teacher explains. Currently, this design includes 5 mini-schools: Nebula which hosts 4 kindergartens, and Orbit, Comet, Galaxy and Eclipse each with 6, 6, 9, and 10 classes respectively servicing first through 6th grade.

Another way the principal has attempted to give the school a neighborhood feeling is through the use of street names and directional signs throughout the building. Students don't just gather in the lobby but arrive at Student Unity Square. And they don't just walk down corridors but stroll down Literature Lane, Bulldog Boulevard, Curiosity Court and Rue de Respect. Huge directional signs remind students of the rules of the road or warn them to slow down or stop before turning sharp corners.

After deciding on a mini-school model the actual transition into teams was more difficult as teachers found themselves having to consider and decide who would work with whom in each mini-school. *“I personally have quite a few horror memories of the process,”* a teacher who went through that process describes. *“That process was very hard. Once you declared what your grouping was and what was important to you, other people’s feelings got hurt or they were forced into groupings where they didn’t feel as welcome or as comfortable. There were a lot of hurt feelings amongst the faculty when that happened.”*

The mini-school structure is designed to help build consistency across grade levels and promote vertical dialogue between teachers working with the same group of students over time. Teachers within a mini-school are also involved with hiring new members if a vacancy occurs. Depending on the mini-school, teachers fundraise together, go on field trips and develop a variety of activities or practices like team-teaching or multi-age classrooms. From an administrative point of view it also allows for the development of a variety of educational programs under one roof. In an ideal situation these would be used to match individual students’ strengths with the appropriate team. The principal explains this philosophy: *“I really believe that every school in America should be kind of a laboratory. What happens a lot of times, and some kids can take it, is that we try to fit every kid to one model when we should be trying to develop models to fit every kid, because that’s why kids start falling by the wayside. They can’t conform to one model, because they don’t learn at the same pace and they don’t learn the same way. But we always try to push them into a mold and it doesn’t work.”* People interested in considering an approach that advocates the accommodation of a variety of instructional environments for children have visited this unique organizational model.

It also provides the administration with a way to look at educational programs and activities comparatively by being able to keep track of student performance by mini-school rather than by classroom. The idea being that how students perform will be a good indicator of how well a group of teachers have prepared them since students are expected to stay in their mini-schools for the range of their K – 6 experience. The principal explains: *“In this model they are basically clumped together so you don’t have kids falling through the cracks. You have a certain group of people who handle those kids, deal with them, get to know them, and can hopefully pace them properly. For research purposes, you could go back and take a look whether staffing, or program, or delivery system and see how effective these people were with a group of kids. So if there was a problem, if in one mini-school the test scores were shooting very high or very low, we could go in there to see what’s going on. Not necessarily that you’re blaming an individual staff member but you’re looking at how that group is working or how the kids are doing.”*

As with most designs, however, there are tradeoffs. While vertical communication has increased and teachers do work with their mini-school colleagues, horizontal grade level communication and between mini-school collaboration is harder to maintain. The

daily schedule makes it difficult for teachers to find the time to meet with same grade level colleagues because they don't share the same planning periods. And while teachers have daily opportunities to meet before or after children arrive and depart, this isn't easy for teachers to do consistently. Teachers state that while they would like to meet more often with colleagues from other mini-schools, they tend to feel the need to have a particular reason to call such a meeting since they know how busy everyone is. What they really miss is not the lack of opportunities for discussion but the lack of daily contact with other teachers who are working to integrate the same shifts in policies and practices that they are. Different grade levels are working to integrate new textbooks, science kits, testing requirements, and policies, and these often can be interpreted or delivered differently by different teachers. Teachers miss the support they receive from same level colleagues around these issues. As this teacher relates: *"I'm sure there are a lot of benefits for the way the building is structured with the mini-schools, but to just be able to pop my head out of the door and say, 'What did you do on that math lesson today? I must have missed something.'"*

Teachers who are fortunate to have grade-level colleagues in their mini-school explain this kind of contact: *"Just somebody to keep you in check like (other teacher) who's my grade partner on my wing. He knows that I'm new to the reading series so he'll just pop in my room every couple of days or every couple of weeks. 'What story are you on now?' and he'll say, 'Oh, you're behind or you're ahead. You need to pick up the pace, or, you can slow down.' And it's just very, very helpful."*

The desire to seek out colleagues for advice and/or sharing of information is not uncommon. But for teachers to express this desire at this time may have less to do with the mini-school structure and more to do with the increasing standardization of the curriculum; itself the result of state standards and testing. Willow teachers are not required to be on the same page by their administration. Nevertheless, one side-effect of assessment-driven reform may be the need for teachers to keep track of where other teachers are in the curriculum or how other teachers are covering the material regardless of differences in educational approaches between mini-schools.

One effect of the mini-school design, however, is that reputations commonly generated for individual teachers within the larger community are also developed at the mini-school level. For example, Galaxy is rumored to be made up mostly of the 'old crowd,' teachers active and involved in the teachers' union, 'the core of the school.' These teachers are said to be the ones 'who know what is going on.' Comet, on the other hand, consists largely of younger, less experienced teachers but who bring together a lot of 'ideas' and do a lot of activities that 'are kid-oriented.' They are also pretty aggressive when it comes to fundraising and don't mind coming back in the evening to run a kid-oriented or fundraising activity. Eclipse brings together a mix of young and old teachers who work cooperatively together and often have new and 'innovative' ideas. They are the ones most likely to 'try something new' or 'attempt multi-age or some other arrangement.' And finally, there is Orbit where teachers run 'very individual, but well-

structured classrooms,’ and yet support and ‘work together fabulously.’ Teachers, for the most part, are extremely proud of the mini-school in which they belong and of the colleagues they work most closely with.

Differences in mini-schools could result, however, in teachers perceiving parents favoring or requesting certain mini-schools over others. In actuality it is difficult to assess whether any reputation would hold true since a regular turnover among teachers means that each mini-school is continuously evolving as new teachers add new ideas to the mix. Furthermore, there are other differences between the mini-schools affecting parents’ perceptions. For example, most of the special education or ‘identified’ students are in integrated or ‘cluster’ classrooms or resource rooms in either Galaxy or Eclipse. Eclipse is also home to an extra kindergarten, a transitional first or Target One, and two first and second grade combined classrooms. And finally, parents often select teachers who they feel best meet the individual needs of their individual children regardless of which mini-school a sibling might have belonged to. For the most part, however, families do remain with one mini-school. This is evidenced in the way the teachers within mini-schools talk knowingly about students and the students’ families regardless of the grade level that student finds him or herself in. This familiarity with the families provides a larger context from which teachers assess and understand children.

Elementary Level State Tests

Elementary students in New York are required to take four state-standardized tests: English language arts (ELA), math and science in 4th grade and social studies in 5th. Parents and teachers agree that the ELA seems to carry the most weight followed closely by math and the least important tests are the science and social studies tests. According to several teachers these value ranks are the result of having the ELA and math scores reported in the newspaper, while the others are not. Furthermore, science is thought to be primarily for program evaluation and is considered less high stakes for students because individual scores are not reported, only the school’s overall levels.

The principal states that all the tests are equally important. However, teachers spend more time preparing for the ELA and math tests than for science and social studies suggesting some form of prioritizing of these subject matters. On the other hand, as one teacher explains, having the principal keep track of all the test scores increases all of their values: *“I get extremely nervous with all of the tests. The principal looks at all the grades. He keeps track of who is doing what. He’s got a list: Ms. X, science this many 3s, this many 2s, you know. It makes a difference if the ELA is in the paper, but to me all the tests are really important because the principal is still looking at the grades that come out of your classroom.”* She also feels strongly about preparing well for the social studies exam because the ELA and math tests are supposedly testing the K – 4 curriculum while the 5th grade social studies test is based mainly on what is taught in 4th grade. *“That’s a serious reflection on me. They don’t learn that stuff in any other grade.”*

The state tests are standards-based and are intended to measure how well students are mastering the skills put forth for each grade level by the state of New York. Teachers know what the tests look like and what standards they are expected to cover. They also know the challenges they face due to such factors as poverty, absenteeism, and individual special needs. For the most part students have done fairly well on these exams and their scores have been steadily increasing. Table 1 describes Willow Valley’s performance on the ELA, math and science tests for the past three years.

Table 1: Willow Valley Elementary scores on New York State 4th grade tests

ELA (all students)	Level 1	Level 2	Level 3	Level 4
	(not passing)		(passing)	
1998 - 1999	8%	49%	44%	0%
1999 - 2000	8%	44%	45%	3%
2000 - 2001	8%	36%	43%	12%
Math (all students)				
1998 - 1999	6%	22%	59%	12%
1999 - 2000	6%	32%	56%	6%
2000 - 2001	1%	23%	56%	20%
Science (all students)	below State Designated Level		above State Designated Level	
1999 – 2000	44%		56%	
2000 - 2001	23%		77%	

The teachers’ experiences, however, preparing, administering and scoring the state tests raise many issues that put into question the link between assessing with standardized tests and student learning. Furthermore, these issues impinge on the teachers’ abilities to work as partners with the state. This section describes some of the issues brought forth through the teachers’ responses and experiences as they prepared and administered the state tests during the 2001 – 2002 school year.

Test scores measure more than student knowledge

Most teachers at Willow Elementary believe that an emphasis on higher standards is a good idea. They make daily efforts to deliver a curriculum with an increase in critical thinking, problem solving and writing tasks. But when it comes to actually taking the tests, they find themselves somewhat powerless in front of certain students who may not put forth the needed effort to perform as well as they can. The principal explains what he does to help ensure that the students take the tests seriously: *“One of the things that we do is send letters home to parents saying kids perform well when they get a good night’s sleep and eat breakfast in the morning. In school I go around the week before each test and meet with every class, and let them know exactly what it is that we’re going to be doing. In some of the suburban districts I don’t think the teachers do anything out of the ordinary, but we really have to try to get the best conditions possible. So I go around and talk to them about the 4th grade ELA and tell them that this is the first time they’re going to compare an apple with an apple. This is the first time every child on this particular day*

is going to sit down and they're all going to have the same test and be exposed to the same thing. This is the first time that they can compare how well we do with how well everybody else does. And we're doing this not just to be compared to other districts but to take a look at our own programs and to take a look at our curriculum. And then we discuss remediation. Basically if you don't do well on this test then you are going to be mandated to take a remedial class and that could be AIS after school. And certainly if you need that then we want to get you that, but if you don't really need it and it's because of your attitude and you didn't take it seriously and blew off the test, then you're still going to get it whether you need it. So I try to be honest with people and really bring it down to their level so they understand what it actually means. And then I go into every classroom for every test and just say good luck, all we're asking for is your best. And then we try to provide the greatest test environment we can."

Test taking is hard work. It involves continuous reading, a lot of writing, a lot of thinking, all the while needing to remain absolutely quiet and still for long periods of time. As a result, the teachers feel that many of the students are not putting in the time and effort necessary to do well. It's not that they can't do it, it's just that they don't do it with the expected level of motivation and persistence. After the ELA, a teacher describes how several students spent 7 minutes on a half an hour assignment and played with their pencils the rest of the time. *"It really bothers me that they could have done better."* No amount of 'read the directions' or 'check your work' seems to alter this fact. Another teacher shares her frustration when she watched a capable student spending too much time filling out bubbles perfectly on the bubble sheet only to run out of time overall. If she could have, she comments, she would have taken that pencil right out of the girl's hand and filled in the bubbles herself.

The teachers agree that the tests are intended to make sure students are learning more critical thinking skills but unless the students put in the effort to work towards these skills this is not going to happen. During regular classroom activities, the teacher is in a position to redirect or introduce new angles of an activity to reinforce critical thinking skills or motivate students to try again. The testing situation expects students to read and respond to a variety of cues and to re-read directions if they find themselves confused. If they are not motivated to try again, to seek an alternative approach to a problem, or to re-read a question to make sure they interpreted the question correctly, and they are not allowed to receive help from teachers, it is not surprising that students will sit and do nothing. As one teacher remarks: *"Some of these kids are not motivated to find a way to solve the problem. Other kids will find a way to find a solution no matter what. I don't know how you instill that in kids to want to. They don't have that drive to want to be correct, to want to solve the problem. I haven't figured that out."*

Testing opens many questions for teachers as they try to make sense of the reasons some students can pass these tests and work until the end while many of the students at Willow can't or won't. Poverty and the associated life styles are often associated with lower level thinking and literacy skills. After the math test, one teacher

comments about this relationship: *“I don’t think it’s the math test. It’s mostly a reading test and some of these kids don’t have the reading skills, they don’t have the thinking skills to figure out what the question is asking. Many of these kids are coming from backgrounds that don’t give them the experiences that give them the ability to think and reason. They watch videos and TV, their choice for free time is not to read and it is not to be read to by a parent. Over the past 4 years with the tests, I’ve noticed that the kids who have those experiences, with families who care about education and read to them, are scoring well on the tests. Home life is such a factor and background experiences play a big role, and no one is paying attention to that.”*

Teachers know and understand that the impact of any home life is much more complex than simply whether children read or are read to, watch too much TV, or engage in multiple cultural experiences. They understand that even children who do succeed in school engage in all these activities as well. In their attempt to understand the associations between home life and school achievement, they find themselves limited, as are most studies on the topic, in truly understanding why students from certain types of homes tend to want to do well while students from other types of home tend not to. For these students, the tests are essentially measuring a desire to read and follow through on the directions, not whether or not they can read and follow through on the directions.

Reading comprehension is the primary skill measured by all state tests

Reading, writing and thinking are closely associated. The tests build upon these three skills in a way teachers feel hurts the performance of students who may not be strong readers or writers. The performance part of the science test is a good example of this problem. The test is set up in the library. There are four tables with five hands-on experiments on each. Tall cardboard walls divide each station so the students can’t see what other students are doing. Every student will complete three 15-minute experiments; either stations one, two and three from a pink test book, or three, four, and five from a green test book. They work alone for two of these and with a partner for station three. To create consistency, one of the fourth grade teachers has been designated to read the directions and proctor the test for all the fourth grade classes. This day, two groups of students will take this test.

The first group is made up of 15 students who require readers and 6 students who work alone. After randomly assigning students to a station, the proctor reads the directions, asks if there are any questions, and begins the 15-minute timer.

Instantaneously the room is filled with a cacophony of voices as the readers read the directions for the different experiments. While they don’t assist the student in actually conducting the experiment, the kind of interaction they offer to students provides a contrast to understanding the experiences of students working alone. For example, the reader doesn’t only read what is written but often paraphrases or interprets what is written. She or he often provides directional clues, for example saying ‘pay attention’

until all the directions are read, or pointing to the question and response area on the page as he or she reads. In contrast the six children working alone appear to read a couple of words, fiddle with the material, go back to reading, back and forth, and may not be reading all the directions or in the right order. The reader can also take the student through the process and pace how much time they should spend on each part, although that is unnecessary since special education students do not have time limits. But the presence of the reader can provide a kind of motivating support or focus that is absent for students working alone. For example, a boy is working with the hanging baskets and has to weigh a container with water. He uses up all the smaller weights but the water is still too heavy.

Reader: *"What are you going to do with the weights?"*

S: *"Take them all out and start over."*

Reader: *"That's brilliant!"*

The student starts over. He measures the water then he starts with a larger weight until he has it even. He counts up the weights and adds them all together.

In contrast, students working alone face many situational barriers that have little to do with how well they understand the task or can carry it out. For example, the directions are often confusing. There is a lot of text on the page so that several students miss the small boxes intended for answers because they aren't clearly labeled as such. Teachers and students seem to understand the pertinence of format over content. In preparation for the afternoon group, a teacher prepares his students for what to expect: *"This part of the science test is in the library. The stations are set up. I will not be up there telling what you have to do. You will have to read all of the instructions, so take your time. But these stations are timed. There will be one station where you will be working in a group, and that will be chosen randomly. I believe you have 15 minutes. Are there any questions?"*

S1: *"Can I get a drink?"*

S2: *"Can we bring our good luck chairs?"*

S3: *"What if we don't finish it in 15 minutes?"*

T: *"Try your best. Know if it is a two-part question."*

S3: *"Is it going to be hard?"*

T: *"This test requires you to do reading so a lot has to do with reading..."*

S2: *"comprehension."*

S1: *"(ironically) oh that has to do a lot about science."*

T: *"Try your best."*

S3: *"Do we get a bad grade if we don't finish that portion?"*

T: *"There are five stations, you have to complete three. Time to go."*

There are other factors besides the way the text is displayed on the page that create contextual challenges to the students trying to get their answers completed in time. Nine-year-olds seem to get stuck on extraneous concerns. Some worry too much about doing the right thing and wait with arm raised to ask if they can pick up a ball that rolled under the table or take too much time closing the zip-lock bag properly when the

directions tell them to. Station three asks students to work in pairs on a ball and ramp assignment, so not only do they have to understand and follow the directions but they also have to quickly negotiate who does what in this make-shift working relationship. Furthermore, the pairs are selected to make sure that each student has been working in a different test book, one green and one pink. What happens, however, is that in these booklets, experiment 3 begins on different pages. As the students read and work on this experiment, they find that one needs to turn a page when the other doesn't or one is on the left page while the other is on the right page. This discrepancy is enough to slow some students up as they question whether it is the same experiment because it doesn't look exactly the same.

Another situational factor affecting the student's ability to carry out the experiment correctly is that the students find it difficult to maneuver their space to both do the experiment and write in their test booklets. For example, one pair of students working on the ball and ramp assignment read the directions in their booklet and then carefully place the ball on the ramp where they think it should go. They check to make sure the cup at the bottom of the ramp is in the correct location and then they release the ball. The ball hits the cup and it begins to slide towards the edge of the table. It should have slid off the table but the girl's test booklet is in the way and it hits that instead and bounces back toward the ramp. When it stops, this is the answer that they measure. This happens a second time. Later on one of the test proctors comes by a different group and warns them to move their test booklet out of the way.

In general, the teachers agree that the science test fairly assesses the content students should know. They are concerned though that reading ability plays too heavy a role in preventing students from showing the extent of their knowledge and understanding. One teacher explains: *"The science test I thought was fair. I think it's a fair test. I think that the problem with the lab part is that if children aren't good readers they get caught in how confusing it is to follow the written directions for that part of the test and it throws them off. And we know our kids know those skills and we know that our kids are able to perform those tasks. But we also know that because some of our kids are not good readers they can't process those directions in the time given in order to do that task well. And I think that was clear when you look at the scoring. All of the identified kids did a beautiful job on the sorting and classifying task because they had everything read to them. Kids who can't read fell apart on that station because they had no idea how to process all that information. I think the state has to look at whether we are assessing reading every time we assess something."*

The tests work against teachers' ability to prepare their students well for them

Teachers are especially frustrated when students who want to perform well can't because the task itself is above the students' developmental level, is presented in a way that confuses students, or was not yet covered in their curriculum. So even though the teachers feel that for the most part the tests are measuring skills their students should

know or learn, they may not necessarily be prepared to have mastered these by the time of the test. Furthermore, there are so many content areas or skills that might appear on any given test, that it is impossible for the teachers to completely prepare their students for every possible scenario. It becomes a gamble then for the teachers to know which skills or content area they should emphasize in preparation for the test. The math test is a good example of these dilemmas.

Up until the day of the test, the teachers express little concern about the math test feeling that generally their curriculum, with some modification of how much to cover before the test, does an adequate job of preparing the students. This year, however, they are in for quite a shock.

At 8:50 on the first day of the math test the principal comes by and tells the students *“Good Luck! I’m sure you’ll take it seriously. You are well prepared. You will have 3 days of this test like the ELA. Good luck!”* A little later he makes the morning announcements over the intercom. He informs the school that the 4th grade math test is today and asks other students to move quietly in the hallways so they can provide the best testing environment for those students.

S: *“I’m nervous.”*

T: *“You don’t need to be, you are very well prepared.”*

S: *“Yes but I’m not very good at math.”*

T: *“You’ll do fine.”*

S: *“Yes but I haven’t gotten the hang of the big questions.”*

S2: *“I’ve got an A in math.”*

S3: *“I’ve got a B+.”*

S: *“I only have one A in my report card.”*

T: *“You need to clear your desks so finish up your breakfasts.”*

The teacher points out several students who *“are acting like 4th graders who are ready to show me their best.”*

While the students finish their breakfast and morning work, the teacher sits at her desk and looks through the test. She turns around with a frustrated look: *“It makes me mad that they constantly ask things 4th graders shouldn’t need to know like percents and probability. Fourth graders should not have to know percents. Fourth graders should not have to know probability. It makes me mad.”* It’s not that teachers don’t teach percents and probabilities or that they don’t believe their students shouldn’t understand these concepts. Percents and probability are taught as early as 2nd grade. But whatever the state expects students to be able to do at this level does not eliminate the reality of what each child needs to know to move ahead successfully in a particular subject. Teachers question whether it is fair or even productive to emphasize mastery of such a wide range of concepts and skills for testing at the 4th grade level. They feel that doing so is likely to negatively affect students’ conceptual understanding of necessary skills such as addition,

subtraction, division and multiplication, which are still in their developing stages of mastery.

A fourth grade teacher explains: *“This year’s math test just seemed to be very heavy on perimeter and area and skills that I don’t think are relevant for 4th graders. We have kids who can’t add and subtract. We have kids who don’t understand how to regroup and how to divide and those are concepts we should be testing. And instead we are asking them to trace pattern pieces onto a piece of paper. I found the directions for a task like the perimeter question and the trace the shape where you had to trace that shape so that it was two times bigger confusing. So that threw off a lot of kids because no longer were they showing math skills it became a reading test again. I just think there were a lot of questions on skills that we don’t spend a tremendous amount of time on and very few questions that really assessed our children’s ability to solve basic word problems or multiple step word problems that are relevant to their life. I mean asking you to figure out a child’s age twelve years from now and then her father’s age based on that piece of data is not a problem that a child is going to encounter in the future. Ask me how much change I’m going to get back after I go to the grocery store and I spend a certain amount of money or how much money have we spent in a day after running errands, those are things children need to be able to solve. So I’d really like to see that test be more relevant to daily living and life skills as opposed to geometry and perimeter and things like that.”*

After reading the directions, the students begin to work on the test. The teacher and proctor walk around, standing over students’ shoulders watching as they work. Occasionally they pat a student on the shoulder or confirm that yes they can write in the booklet. Most students work independently. A couple of students raise their hands repeatedly. A boy who wanted a drink at the last minute but doesn’t get one is restless throughout. A couple of bubble sheets fall to the ground. Overall, however, the students work quietly.

The teacher goes back to her desk and examines the problems on the test. She quietly remarks that some of her students may be stumped by so many geometry questions. She hasn’t completed the geometry unit yet and comments that she can’t possibly predict what to cover based on the test nor does she feel she should change her program around because of it. But as she walks around it is clear that geometry is heavily emphasized this year and the gamble to spend more time on other skill areas like graphing and long division might dramatically affect these students’ performance.

But the issue of what to cover, how to cover it, and whether or not 4th graders can grasp the concepts well enough to respond to them on tests is made more complicated by the test format itself. On day two and three of the test it is obvious that geometry is not only a component of the test this year but has been integrated in other contexts such as graphing. On the third day the teacher who has been walking around and observing the students work, whispers: *“That is the third person who obviously didn’t know that a*

square has four equal sides. The problem asks them to plot the other two corners of a square, but the students don't know that a square has four equal sides. It is very frustrating to know that you've taught something and they act as if they've never seen it in their lives!"

The problem is not that fourth graders do not know that a square has four equal sides, but whether or not they know that this is what the question is asking them to demonstrate. Sometimes the tests use information in ways that are not familiar to either the students or the teachers. The way the question is worded or concepts are integrated into a particular question makes it difficult for some students to know what is being asked and to pull out the essential knowledge in a way that becomes useful for them to apply to the question. Over and over teachers find that the state does not seem to pay adequate attention to how the question is worded and whether or not the concepts included are introducing non-conventional ways of integrating material.

These feelings are not unique to this teacher. *"It is just not what we were expecting,"* another teacher comments. *"I don't know. Maybe it's just us being paranoid, you just never, ever feel like you've prepared the kids enough no matter what. No matter how much you've put out there, when you come to that test you're kind of like shaken up a little thinking I don't think I did that the right way, how it should have been presented."*

On the other hand, this year's ELA, is seen as an improvement over last year according to some teachers. One teacher explains: *"This year the ELA seemed to me a bit more fair than the past two years basically because I thought it was more a reflection of what a 4th grader should be expected to do. Last year they had given picture frames for the kids and they were supposed to compare these two pictures that were in the frames and then write in the empty frames. Well of course kids drew pictures because a lot of the kids don't read the directions or they just get nervous. So I didn't feel like there was as much trickery this year. It just seemed to be geared more towards fourth grade in the way the wording was written on the test. Other years it seemed as though they were setting kids up more to make mistakes right off the bat just from the directions themselves, and this year it just didn't seem to be that way. And I think this year they were able to use more of their creativity, more of their own thoughts and ideas and I don't really think that was available to them in the past two years. Whoever is making this test is finally getting down to the kids' level and you know thinking if I was sitting in 4th grade would I have any idea what this means?"*

Deciding what to teach and to emphasize becomes just as crucial as teaching it well so that the students can recognize it and apply it. These decisions become a source of anxiety and frustration as teachers feel that while they are doing their best to raise the achievement level of their students, the state is playing an unfair guessing game. A fourth grade teacher explains this process: *"The math covers so much material you have to find a way to teach that entire math book by the beginning of May, you cannot leave anything out. And while you're trying to do that you're trying to find time to get in science review*

and prep test for that. And before that the ELA. So maybe you kind of put a few math skills on the back burner and figure you'd have time to pick them up because you're trying to do more ELA prep. And then you know it comes down where it's two weeks before the test and maybe you forgot to fit in perimeter and area somewhere. And if that is what happens then you have to pray that it's not on the test. And if it is on the test then your group is going to do badly just because you, and it is your fault you know, because you weren't able to get your kids those lessons in time. So this year I put division on the back burner which was suggested to me in previous years by some of the teachers who had been here for awhile and luckily there didn't seem to be too much division. And I really pushed things like fractions and perimeter and area and multiplication and using graphs and tables, and creating graphs and tables and giving information and all of those things. Just being able to use manipulatives, just like a ruler and your counters and things like that, that has to go on all year. Those types of things have to find ways to get into almost every lesson that you do when it comes to math so the kids don't take long trying to figure out how to use these counters, because it's a timed test. So that's another thing you have to worry about while you're teaching all these things, is make sure you're using all the tools that they're going to expect the kids to use on the test and to find ways to solve those problems using those given tools."

The curriculum in other subjects presents similar concerns. Needing to cover the curriculum in the core subjects and prepare students for the 4 state tests leaves 4th grade teachers with little extra time to invest in sorting through all the standards and creating innovative connections between subjects. And while they believe that the district is trying to align the curriculum in ways that will benefit student performance on the tests, that may not be enough. There are multiple demands placed on 4th grade teachers and lessons such as those that are involved in the STC kits take time to prepare, set up, cover and clean up; time that does not exist. Furthermore, the kits focus on particular key ideas like floating and sinking or electricity and these often compete with rather than complement the science needed for the test.

A 4th grade teacher expresses her frustration: *"I think our entire science program really needs to be revamped. I think that we need a textbook that is a review of grades one through four. If you can find a textbook that does that it will be perfect because that's exactly what the test does. So we're expected to teach these STC kits, I mean giant lessons with all of these things in them that really are not on the test at all. And the kids are supposed to remember the things that they've done in these kits in the first grade. That's how the test goes. It's unbelievable. Who has a memory like that especially when you're little? Not many people, not many children, I'm sure. So they get to the test and they can't even say the word. They're not even sure how to pronounce the word, you know, chlorophyll or photosynthesis even though maybe they talked about it for three straight months in third grade. It just gets lost somehow. So we have to find our own materials. We go out and buy books from the parent-teacher store. We go on the Internet. We share what we have from past years or just from teachers in lower grades and try to*

create our own science curriculum to get those things out there that they haven't seen in a few years and still do our STC kits."

While the principal and others who have worked closely with the teachers in science disagree that a textbook is a good solution, they agree that science can be an area of frustration because they feel many teachers do not have a strong content knowledge in science. If that is the case then the teachers' frustration with the STC kits might arise from needing to cover too much content for each unit and paradoxically not having enough time to cover all the content necessary, especially in fourth grade. With the help of the Capital Region Science Education Partnership professional development specialist, the district has been working to align these kits with the standards and state tests. And while that may eventually help teachers prioritize the content for science, a concern for teachers is that it seems to them that no one is looking across the entire spectrum of curricular requirements that are not only expected in the 4th grade curriculum but are essential to each particular test, including review.

The stress of preparing for each test creates a form of curricular vacuum where teachers spend so much time emphasizing preparation for each test that when the test is over, there is almost a sense that there is no point or purpose in continuing to teach language arts or math. After the math test is over, one teacher expresses this stress in jest: *"My year is basically over right now. I could do science every day. I could do what I want. There is absolutely no pressure right now. Now I can do things I want to do, the things I feel the kids should know."*

The state's scoring criteria obscures understanding what students know

If the uncertainty around the content and format of the tests aren't confusing enough, the scoring process really makes teachers wonder about the true nature of the state's intent. Scoring the state tests is another responsibility that is placed on teachers and one that brings teachers face to face with the criteria with which they are expected to judge the quality of the students' work. Most state tests have a multiple choice component as well as short and extended answer sections. Willow Valley pays for services provided by their local BOCES (Board of Cooperative Education Services) to score the multiple choice sections while the other sections are scored by teachers in the school building itself.

Four to six teachers work in tight office spaces with piles of tests in front of them. The tests come with a section for teachers to record each student's score and a rubric for interpreting the responses is provided through the state education department. For the math test, a district facilitator comes over to train and monitor the process but for the science test a veteran teacher is assigned to score as well as train the other teachers.

After the facilitator trains the teachers in how to interpret the rubric on the math test, the teachers begin to work quietly on the pile of tests in front of them. They are each

scoring half of someone else's class so that no one either scores an entire test alone or scores their own students. Teachers are asked not to talk unless they have a question or need confirmation or assistance in reading or interpreting a student's response. However, as with most processes, issues do arise. Some issues affect all the teachers, while others only affect individuals. For example, particular questions can hit individual teachers differently and create confusion or insecurity or outrage as to how to proceed. Here I provide a couple of examples of these issues.

One of the questions asks students to show they have understood a data statement by drawing a graph based on questions about the data. One of the teachers (T2) has a question about how to interpret the 3 point rubric. The students were given some information that they had to put into a graph. This student completed all the information except forgot to label one axis. The child has a title, the axis label and the numbers on the vertical axis, the names for each bar (drawn correctly) on the horizontal axis, but no name for the axis itself. He had provided labels for the bars in the graph (i.e., horse), but not the axis (i.e., animals). When the facilitator confirms that the response since it is incomplete gets a 2 and not a 3, teacher 2 is outraged: *"Obviously he knows how to make a graph, why should he be penalized? Does he have a complete understanding of what goes into a bar graph? Yes."*

T1: *"We've had that problem in the past and we had to give them a two. If you feel so strongly about it, do what you want to and give it a 3. If you go by what they say, give it a 2."*

T2: *"I'm right. They are wrong. As a protest I should put a 3."*

The facilitator has the teachers refocus: *"You would like to give it a 3, but give it a 2."*

Teacher 2 reluctantly gives it a 2 and turns to the next student's response where she is faced with the same exact scenario.

T2: *"Two in a row, just like that, penalized. Tell me that kid doesn't understand graphs. Obviously written by someone who's not a teacher."*

She grumbles a little bit on this point but the other teachers are all attending to their own scoring tasks. Reluctantly this student also gets a 2 and the teacher moves on.

The next student is also given a 2 but this time because he or she had all the correct labels. The actual graph, the bars are completely wrong. Now teacher two is annoyed again and speaks her mind: *"This child obviously did not understand the concept of making a graph. But because she was able to follow the directions and knew enough to label she gets the same points as the other who obviously understands how to make a graph but forgets one label. That's not right."*

F: *"That's why you can't compare answer to answer. You have to go by the rubric."*

T2: *"Ok then you can't compare scores. You can compare scores between schools, yet we can't compare scoring to scoring. You're telling me that child has the same comprehension as the other one?"*

This issue came up in different ways throughout the scoring process. For example, a child who gets an answer correct but do not show their work when they are required to gets a zero while a child who shows the beginning of a process but gets the answer wrong gets a one. The teachers have a really hard time understanding how one student should earn a better score than the other. Is the criteria for following directions more important in the state's mind than that of getting the correct answer? Furthermore the issue of determining partial credit has many of the teachers confused and concerned that they are not being consistent throughout. This problem occurs because interpreting the rubric is a subjective factor and teachers understand the criteria they employ in the process of employing it. This means that they are constantly adapting their understanding when faced with new circumstances. What happens is that the teachers become aware that their shift in understanding is a threat to consistency. If they get the chance, they often go back and review certain questions that may not have been scored correctly now that their understanding has shifted. If they don't have time built into the process this search for consistency can be hampered.

After the teachers finish scoring one whole section of the math test and break for lunch, one teacher (T3) is having a difficult time with the consistency issue and feels very strongly that they should all review their scoring especially for the multiple step problems. The facilitator is worried about time. T3 however wants to discuss it as a group *"to understand each other's thought process."* The other teachers agree and so they decide to re-read the rubric and discuss their scoring process.

The issue at stake is how to interpret "beginning of a process." The teachers go back to the rubric and revisit the sample questions on those questions. The rubric states that for a two-point question you may give a one if the answer "may contain an incorrect solution but applies a mathematically appropriate process" or "addresses some elements of a task correctly but may be incomplete or contain some procedural or conceptual flaws." It is the idea of 'some' that is confusing to the teachers. How much of a process is required to get a point? For example during a sample question, the teachers incorrectly give a 0 to a response where the student showed a correct process but gave a wrong answer to part one of the question and did part 2 of the question totally wrong. Teachers are surprised that this response was given a one in the scoring guide because it meant that the student could earn half the points when _ of the response was wrong. As one teacher then interprets it to mean: *"So 2 is completely correct, 0 is completely wrong, so the rest is a 1."* The facilitator says no and has them look back over the rubric but some teachers find it difficult to determine what amount of a response constitutes a 'some.' For example, on some questions the presence of a word or concept needs to be found in the student's response to earn at least a one, while others are seemingly more open to interpretation. As one teacher states: *"The more examples I see, the more confused I am getting."* While they continue to disagree on their interpretations, the teachers do end up agreeing that when in doubt they should err on the side of the student. There are, however, aspects of this that they have a problem with.

T: *"You can have a brilliant child who does all the work in his head correctly and gets all zeros."*

T: *"And I can have a child whose numbers are all incorrect but whose process makes sense and they get partial credit."*

Finally they feel ready to recheck their work. While they do this, there are a lot of quick checks of work with other teachers. Some teachers talk out loud *"ok beginning of a process, it's a one."* Others aren't so sure: *"We're saying that if a child attempts something, that they tried some sort of a process, even though the answer is wrong, he gets a 1?"*

F: *"Yes."*

T1: *"I'm having trouble seeing whether it is a correct process."*

T3: *"All he needs is a start of a correct process to get partial credit."*

T1: *"But in the written response you have to infer what they were going to say to know that."*

F: *"You can't infer what the students meant."*

This concept relies on teacher's ability to attend to the reasoning behind the student's approach. On the positive side, this process gives the teachers an opportunity to discuss some of the ways they are seeing their students tackle particular problems. But it also lays bare the realization that the scores received by students, schools, or districts don't in reality offer any idea as to the students' level of knowledge, skill or understanding. If parts of processes get credit while correct answers with no processes don't; if obvious understanding of graphing gets the same score as no understanding of graphing with correct labeling, then how can scores reflect in any meaningful way a student's understanding and knowledge? Furthermore, the rubric used by the state raises serious questions about what exactly the state is looking for since most of the teachers conclude that, *"none of us as classroom teachers would accept many of these answers."*

When they are all done re-scoring this session, they feel better that they did. They are surprised to find that they didn't have to change too many after all. This desire to be consistent and do the right thing for the students as well as for the state turns what is relayed as a technical process into an ethical, moral and judgmental process that places considerable strain on the teachers to do the job correctly.

Finally perceived discrepancies in scoring criteria only add to the teachers' confusion about the role tests play in assessing students' work rather than assisting them in understanding the state's assessment system. For example, the teachers do not understand how most of the students received 3s and 4s on the social studies exam when that is not the case for the ELA or math tests. One 5th grade teacher comments: *"I have a feeling it had to do with the scoring. Based on their ELA scores, I didn't think we were going to do that wonderfully. But overall in this classroom with 9 special education students, every kid got a 3 or a 4. My sense is that the scoring has to be different."*

Teachers and administrators take the scoring process very seriously and while they may disagree with some of the criteria used by the state, they do work hard to understand and work within the state guidelines. Teachers talk about scoring as being both beneficial to their own understanding, as a time to discuss the rubric with colleagues and understand what the state is looking for, as well as being harmful to their students as it brings a substitute into their classroom that many more days. It is however, unclear, after all the test preparation, administration and scoring, what can really and usefully be learned about the students or the strengths and weaknesses of the curriculum from these scores. When correct answers receive zeros and only partially correct answers receive a one, what kind of conclusion are teachers expected to make about the quality of their instruction?

Teachers are not treated as partners in the state's testing system

Teachers put a lot of effort into understanding the test format, the content of the curriculum, and the scoring criteria. And yet each year, as the examples shown attest, they are faced with new types of problems, a shift in focus, or other unexpected kinds of changes in the state assessments. The effect of leaving the teachers out in the dark on important format, content and scoring questions results in a sort of fatalistic attitude in terms of their own efficacy in preparing their students for the tests. Emphasizing a standards-based assessment approach to teaching and learning while keeping the interpretation of those standards out of the teachers' reach might have the effect of lessening not improving teachers' efforts. While teachers attempt to meet the expectations set forth by the state and their district, the issues raised by the assessment practices may affect the performance of teachers who see their professional integrity tied up in the state's ability to adequately address the validity of the tests as measurements of their students' knowledge. The pressure of turning out higher scores each year regardless of changes in the content and scope of the tests and the individual needs and concerns of a each new set of students does not seem to alter the teachers' dedication to their students. It does, however, seem to alter their enthusiasm for an assessment-based curricular system.

Teachers are experienced in using assessment strategies that look at children as complex and multi-dimensional beings. Keeping an open-mind and a positive attitude about what counts as learning and progress underlies the philosophy and structure of many elementary schools. This school is no different. Teachers have been taught to celebrate individuals' successes and work hard to overcome their weaknesses. They want to see students using strategies they have taught regardless of whether their answer is right or wrong. And they want to see evidence of progress in all areas, not only those measured by state tests. Assessments that work counter to what they have been taught and what they feel is right creates ethical tensions for teachers.

“My idea would be to look at each child and see where they are when they come into my classroom and watch their progress over the course of the year. And at the end of

the year take a look and see how much progress they actually made, from beginning to end. Because for some kids, they may come to me in September and they may not even be able to write a sentence as a fourth grader. And yet I may work with them and work with them, and by the end of the school year, they may be writing a five sentence paragraph which might only score a two on the ELA test. But for the child to go from not being able to write a sentence, to being able to write a five sentence paragraph, that is a world of difference.”

“Instead of that being a negative that they got a two, what a great job they did over the course of the year.”

These issues bring to the surface many interacting components that affect teachers’ work and decisions surrounding the state tests. Students’ personal experiences, how well they are prepared in their individual classrooms, how well they are each able to understand and apply what they are learning, the decisions a teacher makes in terms of what to cover and how to present the information to students, and the decisions made by the district and state at the policy level all affect the students’ abilities to perform on state tests. These multiple and interacting spheres of influence create unique systems; systems that have been in place before the advent of the New York State tests and now continue, albeit in various ways, to coexist with them.

The Educational Climate

I think in order for a school to be successful, everybody top on down has to be on board to what your desired outcome is, and I don’t think they are. Not standing up and preaching to the crowd. Truly living it, being part of kids’ lives, and being part of the curriculum and what is going on in the school building. Right from the top on down needs to be knowledgeable about what kids are learning and how kids learn.

(Willow teacher)

Like many city schools, Willow Valley Elementary has its share of politics, challenges, and resource limitations. It is located in a small city of approximately 11,061 having as its primary employer a federal arsenal established in 1813. Over the past two decades, over 2,000 jobs were lost when several important manufactories, namely in stainless steel and paper, took advantage of NAFTA (North American Free Trade Agreement) and moved their companies out of the country. The effect this has had over time is to transform a fairly stable working class community of Dutch, Irish and Italian descent to one that includes a growing number of people characterized as poor and transient. Parents, teachers and administrators seem to agree that the poverty index level and the transient nature of this population present many challenges. Furthermore, the organizational leadership plays a key role in the overall functioning and climate of a school as does the demographics of the community and the relationship of both together.

School and community climate is a difficult thing to assess, however, and most communities have their share of tensions and disagreements around education. It is not the point of this story to try to pinpoint the causes or sources of tensions but there are a few situations that are important to describe in order to understand the context within which decisions around meeting state standards are occurring. For example, at 9, 219 dollars per-pupil expenditure, money, or the lack of, is a huge topic of discussion at Willow Valley. Furthermore, teacher contract negotiations have been relatively heated and the teachers worked almost this entire school year without a new contract. While the principal is proud of how well they do with limited resources, he is aware of the difference an increase per student expenditure could make for Willow Elementary. He states if you multiply 825 by, for example even 200 dollars, *“that’s a substantial chunk of change that I could use. I can break down class sizes. I can get additional computers. I could get some more one-on-one teacher assistants or whatever we need, and I’m sure we can raise standards even further. Unfortunately that’s not happening.”*

Another source of frustration for teachers, administrators and parents is that many people, top on down, do not seem to be working well together. Besides establishing a strong school vision, the school board makes budgetary decisions, negotiates contracts with employee unions, adopts a school calendar and approves the school curriculum. The superintendent is hired by the school board and acts as secretary to the board as well as the chief administrator for the district. The board also acts as a channel between school and community. How well this group is respected in the district and is seen as responding to parent and teacher concerns will affect how well the district works together to create the best environment for children. At Willow Valley there is no evidence that teachers, administrators, and other personnel are not working hard to create the best learning environment for children. There is however considerable evidence that they are not working synergistically.

Modest pay, contract issues, staff turnover, tensions between teachers and between teachers and administrators all contribute to a system’s ability to collaborate and agree on important issues. Furthermore, teachers report that tensions between teachers and administrators seem magnified when their expectations don’t match. While the principal feels that district and state demands placed on teachers are part of the teachers’ responsibilities and should be met with a strong commitment, teachers feel that the administration does not adequately understand the quantity and type of demands placed on teachers today. Several teachers relate that while they do strive to meet the new requirements placed on them, and they continue to participate in numerous ways on school committees and in school affairs, they feel that little is given in return for their efforts. What is lacking, they explain, is a strong, district-wide, proactive stance that addresses the growing state demands placed on teachers while also addressing the educational needs of the students.

Despite personal and professional tensions and limited resources teachers do stay and work in the district. And they do so with a sense of purpose and hope. Like a large

dysfunctional family people are both interested in making it work and frustrated by the repeated behaviors that don't seem to change. They enjoy working with their colleagues and assisting in the development of new programs to improve the education of the students they serve. But when new blood comes on board there are mixed feelings. On the one hand there is the anticipation of working with someone new and the camaraderie that ensues. On the other hand there is some resentment, as older, more experienced teachers do not always see themselves equally utilized as resources by the administration. *"People in this building are more than willing to bend over backwards to do what is needed, but they need to feel valued and I don't think they are made to feel that in this building."*

Parental involvement

Working with a community stressed by poverty and unemployment is an important contextual feature that cannot be ignored when considering how the state tests have affected this school. The needs of a community are very closely linked to the kinds of services provided by schools. And these services in turn shape a community's perception of the quality of the school. School-community relationships as well as teacher-administrator relationships play a large role in developing the kind of learning community that benefits all those within it.

Parent participation, for example, is an important component in promoting the success of students. Parents from Willow Valley have offered mixed responses about their interactions with school personnel. While some parents report being frustrated at the lack of response to their concerns, others relate having great relationships with individual teachers. Most parents agree, however, that parental involvement at Willow Valley elementary is quite low when it comes to academic issues. They state that while the majority of parents enjoy coming to the school for such events as the annual science fair, grandparent or special person's breakfast, or meet Santa, they are much less likely to turn up for the PTA or other meetings around academic issues. Several parents comment about parental participation at Willow Elementary:

"I think mostly parents should be more involved. I mean you have a city which has at least 10,000 people in it give or take a few thousand and we are like the only few parents here. That's sad, that's really sad. They can be out there at the arsenal for the Christmas tree lighting, so why can't they be in here for their child's education?"

"Our son, the youngest, is the fourth one to go through this school district. And there have been many years when I wished that I had a crystal ball and could see what the future is going to hold. They've all done extremely well. So for us you know we'll stay because I pay school taxes and I will be dipped if they are going to run me out. I will sit on every committee that I can and I will aggravate the school board and I will aggravate the principal and you know if they pay me lip service, well you know I pay their salary. And all I can say to the parents that are new to the district, keep plugging, keep going to

the meetings, keep making your voice heard because we've seen changes and when parents band together you can make a difference."

"But, there's an awful lot of apathy in this district."

There are other issues that affect the day to day teaching and learning that goes on at Willow Valley. Some of these are beyond the control of teachers and administrators, while others are felt to have evolved in ways that work against the development of a strong learning community. Two such issues came up frequently in conversations with staff: the actual instructional time received by students and the emphasis on competition as a motivational strategy for teachers and students.

Instructional time

Children spend a state-mandated amount of days and hours at school. What is done during those hours, however, can vary from district to district and from school to school. Schools make a variety of decisions that affect what students do during their hours at school. At Willow Valley, for example, students have breakfast from 8:35 to 9:05 in their classroom and then dismissal for buses starts at 2:30, as there are 11 buses to load. Add lunch and the specials and you have 4 hours and 10 minutes a day for instruction.

The concern is not that students are not getting an adequate education or that programs, such as the breakfast program, are not important. The issue is that meeting the multiple needs of children creates tension for teachers whose primary task is to instruct the students the curriculum.

Most teachers do use breakfast for morning routine assignments, either math problems or Daily Language Activities (DLA). And while teachers expect the morning assignment to get done they also seem to treat this time as more informal, and use it for planning with an assistant or to work one on one with students on problems the student missed or did not yet understand. And students respond accordingly and tend to see this as a time to eat, catch up with friends, or organize backpacks or desks, while focusing occasionally on the required work. Teachers do complain about how short the day is when you count in breakfast, the specials and lunch, but it is uncertain whether or not they would feel compelled to change that early morning routine. There may be other psychosocial benefits to students in keeping it this way.

Another area of frustration for teachers is handling effectively the multiple interruptions caused by the continuous comings and goings of students receiving or participating in remedial or other services. For example, one non-integrated classroom of 18 students, has 6 students who get reading 3 times a week, 2 students who get math 3 times a week, 2 students who go to the resource room 3 times a week, 1 ESL student who gets ESL services 3 times a week, and 2 students who go to band 3 times a week.

Teachers, administrators, and remedial teachers disagree as to the best approach for providing remedial services to students who need them. New curricular demands by state and district officials only serve to increase the conflict most teachers feel as they simultaneously want less interruptions or disturbances in their classroom but seek increased services for students who continue to lag behind. Many teachers feel that students, especially those struggling with reading, would benefit greatly from daily remedial interventions, rather than the three half-an-hour sessions students are currently offered.

Working with students with a variety of needs is not new to teachers in this school since Willow Valley has been employing an integrated model with special education students for over a dozen years. This model, which has been visited by schools interested in inclusive approaches, involves integrating special education students in regular classrooms and having a special education teacher working closely with the classroom teacher. The Director of Programs and Services explains: *“In our district we started integration with the elementary school 14 years ago. We call them clusters. What happens there is the special education teacher will go in and work with the regular education teacher and when the special education teacher leaves, the teacher assistant will go in and work with the students. The trend has been to try and make about a third of those students in those particular classes special education students. So if you had a class of 24 students, ideally there would be eight special education students in the integrated classroom. I think the effect has been very positive because not only does it help them academically but socially as well with factors that in the past they weren’t part of. I think it’s been definitely positive for both regular and special education students because when they leave school and go out in society that integration is going to follow them. But I think the biggest impact with the integration is that it’s helped with the state assessments because now they’re teaching to the integrated group and special education students have got to take the state assessments.”*

Remedial services are expected to follow this tradition of integration by providing push-in services to regular classroom teachers within their classrooms. This doesn’t always occur, as the principal explains: *“There are remedial reading teachers who choose to pull their students out, that’s their choice. We would like with all our services as much push-in and integration as possible. A lot of times that is possible, sometimes it isn’t.”*

But again as in the morning work, these interruptions are not always negative and can contribute to a sense of family within each mini-school. For example, teachers often share students for special projects or they share time-out responsibilities between grade levels so that a student from another class might walk in to serve their time-out. In those cases, the teacher receiving the students usually draws them close, asks them what they did, pats them on the back and has them sit somewhere quietly. Children also use other classrooms to finish work or quizzes as this little boy coming into a 4th grade class very

seriously explains to the teacher, *"I'm not in trouble. I just got to do this,"* pointing to the worksheet in his hand.

Competition

Another issue that came up in conversations with teachers is that the principal encourages competition between individuals and mini-schools as a way of motivating people to do their best. A teacher tries to explain his perspective on this: *"I think the principal has always been an idea's man; I want to be different from the next guy or the next school. And he works hard to do that."* She goes on to state that among the different ideas he has implemented at the school, *"I think he believes in competition and that competition is a positive force."*

For the most part the teachers feel that the emphasis on competition, especially the promotion of competitive comparisons between mini-schools has not worked well for them or the school. A teacher explains: *"Initially the administration took a management style of competition between the mini-schools. He would basically try to motivate you by saying 'see what this mini-school did?' That really worked against him as far as morale went. And we reflected to him frequently, 'you know that's not the way to motivate people, you don't feel like doing something when all you hear is school X did this and that.'"*

Several teachers report that while they don't mind hearing what other mini-schools are doing, they do mind having their work compared to the work done in other mini-schools or having their students' performance compared to the performance of other students for motivational purposes. Not only do they feel that this works against collaboration between mini-schools but it also goes against the foundational premise upon which the mini-school model is built. The frustration for teachers comes from having to negotiate the mixed message they receive from administration. On the one hand they are encouraged to develop their own unique mini-school approach and philosophy, while on the other, feel that they are being pushed closer and closer together to reproduce programs and practices that are closely aligned to the state tests. This tension has been further exacerbated by the way the test scores decontextualize the practice of student evaluation from the unique instructional context within which it occurs. The principal feels, however, that teachers don't always understand the public nature of public school and that they can't just sit quietly in the shadows away from public scrutiny. He states that while schools currently have a captive audience, the future of public school is in danger and teachers and administrators will need to compete more than ever with other educational institutions for public approval.

The principal comments: *"I think the testing has certainly made teachers aware that they have to be accountable to a public, because you're sitting out there on an island. If you don't do some things and everyone else is doing them you stand out a little more. The way most schools are designed you have a few people running with the ball*

and everybody else will sit back and do nothing. In this system you can't do that as much, you know, people know what Comet is doing and what Orbit is doing. They know if they are doing activities with the kids. But competition could be a bad thing too. When it gets real competitive people could just teach for the test, they could cheat, they could do all kinds of things, take statistics and blow them up. So I'm not saying that all competition is totally great either. I think a little bit of it is good because I think people get too complacent. I think Michael Fullan said, and I really believe this, that if you are content you have a tendency to get a little complacent and not want to change. What you have to have is your building, not in turmoil, but you have to keep them a little off balance or else they're never going to move. Because if everybody is totally happy and totally content, you know they're just going to stay like that and then you're going to get behind."

The arrival of the state tests creates a new measure of comparison for the principal to use. As one teacher explains: *"He's a former coach. He keeps score. That's what they do, keep score."* In fact the first year of the ELA, teachers describe how horrified they were when he pulled out a chart during a faculty meeting listing all the 1s, 2s, 3s, and 4s received per 4th grade classroom on the ELA. The teachers hadn't previously had a chance to review these privately and felt unprepared and embarrassed by this exposure. And they couldn't help feeling as if they were being compared to each other even though they each had very different students.

The principal explains his perspective on this practice: *"Every year I break the tests down by mini-schools. The teachers were very uncomfortable the first time I put that out and I only did it as an in-house thing. Technically in other districts they have multiple schools so when (another district) was putting up (different elementary schools) we're just putting up Galaxy and Comet and it's just kind of a microcosm of the bigger picture. They feel it's more like here I'm putting up Mrs. Brown and Mrs. Jones. We've been doing this for three or four years and there aren't any patterns that have developed that are red flag types of things at this point in time."*

As teachers face new demands as a result of state standards and testing, new pressures to compete against each other arise. They feel that the test scores replace other, more varied forms of assessment that could be used collaboratively rather than competitively. But as one teacher explains: *"We try not to do that to each other. The competition gets put on us from administration and we try not to let it happen. That's as far as it goes."* Instead teachers share materials, lesson plans, and provide teaching tips and support for each other whenever they can.

The Trickle Down Effect of Testing

I think the administrators are feeling the pressure. You know the building level administrators feel the pressure from superintendents. The superintendents feel the pressure from the board of education. The board of education feels the pressure from the state, and it just

trickles down. And I think that those pressures eventually end up down on the kids, where they don't belong. And it's because of the way those scores are reported. If the scores weren't reported in the newspaper, I think you'd be able to use them differently, because I think what this test is testing is good. I think that kids should be able to read a passage and respond to it in writing. That's basically what the test is asking kids to do, and I don't think there's anything wrong with that. What's wrong is the way the adults in the world take the scores and report them. (Willow teacher)

All organizations cultivate particular cultures that evolve and change over time with the coming and going of individuals and the advent and dissolution of ideas and programs. And so part of observing how this school has responded to the new state-mandated curricula and tests, is to try to understand how it works as a system, and how, if at all that system has been altered or reorganized as a result of these state mandates. The advent of the New York standards and tests has brought about many changes that have either magnified or altered Willow Valley's practices. Several of the challenges and tensions described in the previous section have been altered, improved or exacerbated by the introduction of more rigorous standards and tests at the state level and increased consequences for schools and districts. And many of the initiatives have opened up dialogue around important disagreements.

Altering practice

In search of a blueprint

Probably the most commonly practiced district and school-wide response to the New York State standards and tests is the adoption of new curricular materials. As other districts have done, this district has reviewed and adopted textbooks in three core areas, reading, math, and social studies, with STC kits forming a consistent base for science. The principal explains: *“With the high stake's testing I think we've had to try to keep up and try to educate our parents and our board on trying to make sure that we have resources to put into the hands of our teachers that are going to be effective with kids. In the last three years we've got all new textbooks. When I first came here they didn't have a curriculum and I thought that was too risky. The state tests have kind of forced people to really take a look at the standards and the curriculum and make sure that there are not a lot of gaps in it.”*

Each curriculum is developed and textbooks selected because of how well it aligns with the standards. This approach is also meant to share the responsibility of educating students across grade levels as everyone is working from the same series, and each year supposedly prepares students for the next. And while the classroom practices at Willow Valley Elementary exhibit a wide variety of teaching styles and learning configurations, there is evidence that having a standard curriculum across grade levels

has provided some standard practices in the classroom. Figure 1 provides a list of the series adopted by the Willow Valley school district.

Literature Works. (2000). Silver Burdett Ginn. Needham, MA.

Scotts Foresman-Addison Wesley Math(1998). Mento, CA: Glenview, IL.

New York: Adventures in Time and Place. (1998). MacMillan/McGraw-Hill. New York, NY.

In science there is 1 STC (Science and Technology for Children) kit per year in kindergarten, two in first grade, and 3 STC kits in grades 2-6, developed with the support of the Smithsonian and the National Science Foundation (introduced 1998).

Figure 1: Series adopted by Willow Valley School District

For the most part teachers are supportive of the idea of a standardized curriculum and hope these efforts will pay off in students acquiring the skills they need to be successful in school and beyond.

“I think the major advantage of using the series over what was done before, the same skills may have been taught but not necessarily with the same focus. So you would have students coming from different 2nd grades into different 3rd grades and maybe the focus was on a different act of comprehension or critical thinking. Not that they weren't exposed to it, but this way it's a more consistent thing. So that when students are regrouped they've at least been exposed to the same focus and I think that the structure of the math series is a lot more critically in depth than what has been done in the past. The requirements, having to be able to explain not just give your answer, but the process. And being able to describe the process as a 1st grader is very different from describing the process as a 4th grader. And in the past we wouldn't have expected to describe. So that's some of the changes and the benefits, but I really do believe that the series that we've chosen recently was driven by what is expected on the NYS tests.”

“Oh yeah, the math trying to explain it, explain it, explain it. And as much as we hated the series the first couple of years, I do have to say...”

“It definitely supports what we need them to do. Very much so.”

“The part where you're stressing more of the process now, how did you get there? What steps did you need to get there? I think for me that's more important than just saying well what's the right answer. Ok, yup 36 move on. Because you know then you're reinforcing what the kids know and they're also getting a chance to say how they got their answer and it's also showing other kids that there is more than one right way to

come up with the same answer. I think that the tests definitely allow us to reinforce that with our students. That nobody has just the one right way."

The teachers do not feel, however, that a standardized curriculum is without problems. While it pushes teachers to meet the standards, trying to meet those standards may actually limit the kinds of experiences teachers are providing to students. And they are not sure whether that is entirely a good thing.

"I'm seeing that there is more pressure to meet standards. And it leaves less time to do a lot of the real hands-on type of things, like the extra activities that you would like to do to reinforce their learning. You're really always thinking I have to keep right to the curriculum. I have to meet standards. I have to have these kids ready."

"I think my concern is not so much that they don't have the skills but that they haven't applied them consistently enough to really internalize them. They have been exposed and maybe at the time that you were checking for mastery the mastery was there but the true level of mastery may not be there because they don't retain it or they departmentalize their learning. They learn the math as it applies to math, but it has nothing to do with science. Even though they're constantly showing them this is what we did in math, now we're going to go and we're going to measure this and we're going to take readings and how much has the level changed in the beaker? They don't view that, at least here they don't say, "Oh, wow, this is what we did in math!" Even though you've told them or the next day you go over it, it's like you see some of them looking like, "Did we really do that?" "Did I hear this before?" So there's a level of attention. I think that so much is thrown at them that some times they don't get a chance to really live with all that they're getting and make it their own."

"I think the positive effects would be more greatly felt if we had more time in the day. I don't find that school is as much fun for fourth graders as it used to be. You don't take the time to do the expanded projects that really get them to buy into or apply what they are learning because there is just so much crammed into the day. It's difficult to keep learning as free as it should be when you have the kind of requirement that is there, so yes it has some advantages but I think it is at the cost of children being children."

"And in my position as a kindergarten teacher, I feel that there are more and more expectations of five-year-olds. We're kind of getting away from the developmentally appropriate activities and getting really into reading words and basic addition. And that never was part of kindergarten."

And parents have noticed these changes as well:

"It seems like they spend an awful lot of time preparing for these tests. It seems like their regular, normal education gets lost. Our son took the social studies test this fall and it seemed like an awful lot of pressure on the teachers in that they have to spend a lot

of time gearing up for a test that they don't even get the results till after the child is gone. And it seems like they don't get to teach what they're supposed to be teaching."

"My daughter is in 3rd grade and I'm blown away at what they're teaching these kids. I mean they're way beyond what we learned at that age level. It's just on a very high level, fast paced for the kids to keep up with."

Teachers agree. They comment that while they think there are some advantages to working from a reading series that keeps everybody on the same page so to speak, the series itself has some drawbacks. Some teachers have mentioned that the reading level is simply too advanced to be used adequately with a group of students whose reading levels vary so much. Others have commented that the series does not serve whole language processes well by teaching skills in isolated ways and not providing selections that apply the skills they are learning. As this teacher comments: *"What I end up doing is I end up double teaching because I'm teaching using the series and I'm also teaching using the strategies and the plans that I had when I taught the novels and I'm basically double dipping for them. But you have to in order for them to get all of the skills. And I can't teach skills in isolation. What good is teaching them the 'AT' sound in ten words if they are not going to use it within a story and be able to read it. So then you look for stories like Little Bear and different kinds of stories like that that would have that short 'A' sound within it, so they can apply the skill they learned."*

Finally, the textbooks promote integration across subject matters but none of them integrate with each other. As one teacher asks, *"you have to wonder do you do the math in the reading series or the reading in the math series?"*

Professional development

Professional development workshops are, of course, not new. Different program initiatives have been implemented over the past two decades, which may have positively affected the educational outcome and services offered at Willow Elementary. Some of these have little impact on what is done in classrooms, while others seem to have more. For example, the dropout rate went from 7.1% in 1999 to 2.0% in 2001, putting it below the overall NY state rate of 3.8%; possibly the result of a district-wide drop-out prevention program.

Professional development, however, also reflects changes in mandates and policies coming from district, state and national levels. The curriculum coordinator explains that there are many new and exciting initiatives occurring in the district, from workshops that look at particular instructional activities like differentiated instruction and behavior issues, to district-wide efforts such as strengthening literacy in families and in preschoolers. She states that the *"superintendent is committed to really offering as much as we can in terms of resources for training, workshops or anything that people feel are necessary."*

A main focus, however, of newly developed workshops is the close analysis of the curriculum to make sure it is well aligned with the state standards. For the past several years, Willow Valley, as have other districts, has been involved in district-wide development and planning in response to the state requirements imposed on them. For example, a teacher who works on the CDEP (Comprehensive District Education Plan) committee states that it is a plan that looks at the whole district in regards to strengths and weaknesses, especially in terms of meeting the state standards and passing the state tests.

So a link is formed between what the tests ask of students, how the students perform on the tests, and what the district, teachers and administrators can learn from all of this to inform reforming the curriculum. The principal explains: *“I think the data we bring out to show where we need some areas of improvement is starting to drive some professional development. And with that professional development I think some of our teachers are gaining some skills that they need for the kids. When you are in a small city, an inner city, and you have a high at-risk population, a poor population, it’s easy to get down and kind of give up. So the thing is how do we overcome that? I mean how do we keep everybody moving in the direction that we want them to move, making strides you know with the standards, and making our kids competitive and giving them a good education? The standards I think have brought that throughout New York. It’s had everybody at least reflect on their practice which I think is a great thing. And it keeps everybody’s foot to the fire a little bit which is good too I think.”*

The curriculum coordinator describes such an initiative in science: *“In the professional development plan for next year, parents, teachers, administrators, students, community members have gone through and really looked at the programs within the district in an effort to improve them and one of the areas they’ve identified is the longitudinal alignment of the elementary science curriculum. So this group will focus specifically on the STC kits, how are the kits being utilized, and what do we need to supplement the kits in order to meet the expectations at each grade level? That is a common theme that I’ve been hearing from teachers that they cannot rely solely on the kits in order to get the job done in terms of what the state is saying we need to cover. And after looking at the tests and figuring out what students are being tested on, teachers are constantly asking for supplemental books or a textbook to supplement what they’re doing. So I think the initiative in the district is really trying to identify people who would really like more support and making sure that they have it. And also looking at the documents that guide what we’re supposed to be doing every day in the classroom and making sure that it is in line with what the state is saying, in that it’s getting kids prepared to do well on the assessment. Ultimately those are the measures that we are going to use to assess kids and so what grew out of this I think is that we need to do more work with looking longitudinally at the curriculum.”*

She goes on to explain that one strength of this district is that it is relatively small and teachers and administrators know the children and families they are serving. So this is a place where children can feel connected to faculty and faculty can work with the

individual needs of children. And because the student-for-all plan really applies to all students, the nature of teacher professional development has taken on a broader scope, training regular teachers in understanding special education or remedial issues, or training remedial teachers in standardized assessment practices. The Director for Programs and Services comments: *“Every student has to be tested. And because of that there has been more emphasis on staff development in our district in the last few years to provide workshops, training for the teachers, for both special ed. and regular ed. on how they’re going to work with both sets of students in the classroom.”*

Knowing the tests

According to Willow teachers the best professional development offered in this age of standards and testing is exposure to the tests themselves. They acknowledge that seeing the tests have refocused what and how they teach. And these changes have not only occurred for 4th grade teachers but have trickled down to other grade levels as well.

“I have three sets of tests and it’s really important for a teacher to know, and not just a fourth grade teacher, but it’s really important for every teacher to know exactly what kinds of questions are asked, how the questions are asked, the format of the test, and the time limits set on the test. Those are all really important things for teachers to know specific to the test. Not just what the standards are, but how are kids scored on the test.”

“It has definitely changed the way I teach reading and writing and that didn’t happen until I went and scored the ELA for the first time.”

“I got a chance to score the ELA myself and what I brought back to teaching the 1st grade is when I call upon the students I make sure that they have to answer in that complete sentence. Because a big thing on the ELA is taking that question and putting it into a complete sentence, you know, taking part of that question and using it. So I use that a lot in my class.”

This practice is not limited only to the state tests. Teachers learn to adapt the curriculum to match other assessments students need to take. A second grade teacher explains that the first time her students took the Terra Nova they didn’t do that well. After looking over the test, she determined that it wasn’t because her students didn’t know the material or weren’t able to do it, it was mostly that they had not been exposed to it in the same way. For example, she used to give her students sentences with mistakes in them that the students were expected to correct and rewrite. On the Terra Nova they would give a sentence with mistakes but would ask the students to identify the area, section one, two, or three where there was a mistake. It’s not that the information was any different but because it was presented differently the students got confused. Similarly in math while she used the term ‘math sentence’ the test would call it ‘word equation.’ So now she makes sure to teach both terms in her math lessons.

The push that we needed

Teachers and administrators agree that while there are aspects of an assessment-driven reform that they don't like, several positive initiatives have occurred in their district as a result of the state standards and tests. For example, teachers have expressed feeling more motivated to perform in ways they may not have been consistently doing before the tests.

"I think the standards have opened our eyes to a whole realm of different things that we were never really considering before. And the way that the tests are structured has changed the way that we present information to kids."

"I think it has brought up a level of awareness towards what we need to do, the pressure is actually, we don't like it, but it helped."

"I've been teaching for over twenty years and I haven't changed a heck of a lot other than the content. But I did change the extent of the writing that I ask of my students. I would call it more prescribed. No matter how much I tried to develop my teaching of writing skills before, it wasn't until these tests that I am personally pleased with myself and what I am doing with writing. I took one of CRSEP's courses this summer and I don't think I would have done that before the standards. I really don't."

They worry that without the tests while at first they might rejoice, the result would be negative on the quality of their teaching and therefore the students' learning. *"I would worry that without the tests there would be this ahhhhh! And relax on the quality and the drive. I think they have been good to push us. I would worry that it would be a very negative effect if it wasn't there to drive us."*

Teaching the curriculum (with the tests in mind)

Most of the teachers are pleased that the district has made some attempt to synchronize the curriculum because this consistency is not something they have had much of before. They also feel that even though they are expected to use a common set of curricular material they still have considerable flexibility within their own classrooms and feel supported by the administration in their abilities to try new things. *"You do have a lot of decision-making in your own classroom, how you want to set it up, how you want to run your day."*

The principal agrees: *"I don't think the state standards standardized what's happening. I think that they expect certain things but I think that there are ways to get to those. I think that they give you enough flexibility that you can be creative about how you get there. And you know a standard kind of repeats itself through the grades so it depends a lot of times how the individual interprets it in that particular grade and how much weight that they give to it. I think the standards are trying to weave a connection through*

it all. So I'm not asking them to go lock step in their textbooks. The textbook is your backbone, your footprint. You make sure that the basics are covered. If you want to skip this chapter because you know the kids have it or skip five pages, you're the professional, you're with them every day, if that's what you've got to do that's what you do."

So it is important to point out that even while creating standard practices across grade levels and classrooms, there are still a variety of teaching and instructional methods that occur across the school. It is not unusual to see the children in a variety of configurations such as in small group work, peer editing, having older children assisting younger children, or children reading or working in the hallways. Teachers also use a variety of multimedia. While not heavy on the use of technology, they do incorporate drawing, collages, cartoons, model making, and other hands-on work within the curriculum. And teachers do find time to do activities that they have always enjoyed, regardless of the state tests. For example, one May morning in several of the third grade classrooms you could hear the quacking of ducklings as teachers borrow eggs from local farms, incubate and hatch them in the classroom sending them back after the children have had a chance to witness and experience their birth and growth.

The tests, however, have changed the priorities placed within their curriculum and the focus of their lessons. Because of the tests, teachers are making a conscious effort to follow through and make sure their students explain in math or find the textual evidence in social studies, instead of just providing an answer. A teacher explains:

"The tests have forced me to do more modeling with my students than I used to do in the past. I would model and then expect them to get it and now I realize it just doesn't happen that way. There's a tremendous amount of modeling going on and if you want the finished product to have a certain criteria, I model almost every step all the way through. And you know gradually wean them to that level of independence."

These decisions often create tension, as this teacher explains: *"I really haven't seen the tests. I saw the social studies test. And I saw some of the questions on that, and I'm thinking, 'is that really important for kids to know?' But the English test, from what I've heard, seems like kids have to be able to read and extract information from it, or know what the main idea of it was. They have to be able to express their thoughts in writing. I don't know the details but I think that's good. The math test requires kids to write more, requires them to not just solve. They have to read word problems. You may know how to multiply or divide fractions, if you don't know when to use that, that's not good. On the other side, if you can do all these word problems, but don't know your basic math facts, which the kids don't, you're in a catch-22. You want them to do this critical thinking, but don't have time to spend on the basic math facts, then kids can't add, subtract seven from nine in the fourth grade."*

One of the ways teaching seems to have been altered is how assignments are viewed and taught. For example, there is a lot more stress on an outside audience.

Students are no longer encouraged simply to do their best but to consider what they are being asked to prove to a non-existent audience. There are also many examples about ‘talking about writing.’ The emphasis is on talking about writing and reading, not about experiencing the story itself.

In one fourth grade they are reading a selection from the reader about endangered species. The teacher shares a story about some people who chained themselves to a tree to stop them from chopping it down. Some of the kids are impressed. Others comment that that’s pretty stupid. Then they read about manatees. At one point the teacher stops to go over some of the vocabulary they’ve been encountering.

T: *“What does on the air mean?”*

S: *“When it’s on the air that means it’s on TV and not off.”*

T: *“What does sponsor mean?”*

S: *“It gives them money to help keep something on TV.”*

T: *“Someone says that manatees are going off the earth because they don’t have a sponsor. What does that mean?”*

S: *“Like it doesn’t have a commercial for it.”*

T: *“You’re on the right track. Look at the tiger and the manatee and give me your opinion about them.”*

S: *“The tiger looks cuter than the manatee.”*

S: *“The tiger has beautiful fur.”*

T: *“If he had said that tigers have fur, that’s a fact. When he says beautiful, that’s an opinion. Back to my question. People like tigers better than manatees. People sponsor what they like more.”*

After they finish reading the story the teacher asks the students why they think the author wrote the story.

S: *“She wrote the story to encourage you not to hurt the animals.”*

The result is that students are heard talking about the process as well as the product. For example, in a third grade class, the students are supposed to be writing a story about a special person. The teacher is in the back of the room typing up a student’s finished story. A boy and a girl are at their desks and begin to pick on each other. The girl points out to the boy that he should be working. The boy comments that she should talk since she hasn’t even begun. He at least has a whole paragraph written. To which she answers spitefully: *“Do you even know how to start a paragraph? Do you know that indenting starts a paragraph and upper case?”* He answers as a counter-attack that a paragraph has 5 sentences. Ammunition indeed!

Similarly in math, teachers do a lot more modeling and problem solving as ways to talk about the assignment. An important component of these lessons is to incorporate the correct vocabulary; the kind students will see on the state tests and need to use to answer the questions. In a fourth grade cluster class, the special education teacher (ST) is giving the math lesson. It is on polygons. On the overhead she has written ‘Plane figures – flat. Polygons – closed (flat) figure made of line segments.

ST: *"Would a circle be a polygon?"*

S: *"No."*

ST: *"Why not?"*

S: *"It doesn't have line segments."*

She draws two line segments.

ST: *"Two line segments, can I make a closed plane figure?"*

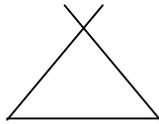
S: *"No."*

ST: *"So I probably can't with 3, huh? If you think you can make a polygon with three line segments raise your hands."*

Many hands are raised.

ST: *"It can't have any gaps. It has to be closed."*

She puts 3 pens on the overhead and has a student come and make his figure. He creates this:



ST: *"Did you close up your polygon?"*

S: *"Yes."*

ST: *"Did you use line segments?"*

S: *"Yes."*

ST: *"Did your line segments meet at the end?"*

S: *"No."*

ST: *"What can you do to have them meet?"*

He fixes it.

ST: *"What is this called?"*

S: *"Triangle."*

ST: *"'Poly' means many and 'gon' means sides so a polygon is a many sided figure.*

There are many kinds. It has to be closed. And it has to have angles where the lines meet. Now there's a new word."

The teacher from the back of the room interrupts and says: *"Boy I'm having a tough time figuring out tomorrow's bonus spelling word."*

S (in chorus): *"Angles."*

Other ways that teachers have changed the way they deliver their curriculum is in how much reading and writing they now do across the subjects. *"The amount of time that I give to social studies and science hasn't changed, but the writing is a much, much larger component. I'm looking at the social studies test that I gave a few years ago, and the majority of it was just spit that content to me and maybe a little essay here or two that was a spit back. But now seventy-five percent of it is written response that you've got to take knowledge from, you know, go back three units ago and remember what we learned about this and compare and contrast it, which I would never have asked of the children even four years ago. And they can do it. It always amazes me. Look! They got it!"*

Teaching to the test

The problem with the numbers [test score reporting] is that it seems like a lot of what we do is test-driven. You know they tell you you're not supposed to teach to the test, but in effect that's what ends up happening. You have to teach to the test, because if these scores are not good enough, then it's reflected back on you. It's reflected back on the students. It's reflected back on the administration, and that's NOT a good thing.

(Willow teacher)

The teachers at Willow Valley may select different test preparation books to prepare their students for the test, but for the most part the test preparation activities conducted in the classroom are very similar. Teachers know that part of their task is to prepare the student for both the test content and for how to understand what the test is asking them to do. Most of the teachers use old tests as the main part of test prep. *“One thing I did do that was very valuable is using the old tests to prep them, that's valuable to see how they ask the questions, what kinds of questions will be on the test, what the test will look like.”* Preparation for the ELA provided many examples of what this looks like.

Excerpt one

It is December. Today in this fourth grade class the students are finishing up making paper wreaths and writing journal entries based on this prompt: “Pretend you are a friend of Rudolph's and the other reindeer start to make fun of him. What do you do?” Under the journal entry are two sentences listed for DLA (Daily Language Activity):

1. Do you want to ride in my sleigh asked Santa
2. I answered yes I'd love to

The students work on these assignments while they eat breakfast. Then the teacher calls everyone to morning meeting on the floor. When they are all settled, she writes on a white board ‘its’ and ‘it's’

its – this it shows possession – someone owns something
it's – this one is a contraction for the two words it is

T: *“So when you are trying to get the 3's and 4's that you are all capable of, you need to be careful of these small details. And also check for this. I see a lot of, ‘I walked in the woods and saw a rabbit and it was hurt so I got my mom and she was scared to help it.’”*

She goes over this sentence with the students and helps them break in into smaller sentences.

T: *“Get the grade that you deserve guys, edit your paper.”*

She tells them to always find a peer to edit or look over their work. Then she passes out an ELA worksheet and sends the students back to their desks.

T: *“There’s an extra page, what are you supposed to do on your planning page?”*

S: *“You’re supposed to organize your ideas so you can take them to write your answer.”*

The teacher points out how one student used a web but you could also make a list to organize your ideas.

T: *“Do you get graded for your work on the planning page?”*

S: *“No.”*

T: *“Are you going to spend a lot of time on this page?”*

S: *“No.”*

T: *“Why not?”*

S: *“Because.”*

T: *“Because it’s a timed test. You only have I think 45 minutes.”*

The teacher then has a student read the directions on the page telling them to read the whole question. She has them underline that statement.

T: *“Are you going to get a good grade if you only answer kind of the whole question?”*

S: *“No.”*

T: *“Maybe the whole question?”*

S: *“No.”*

T: *“You need to answer the whole question.”*

(...)

T: *“This passage asks you to compare. What does it mean to compare?”*

S: *“It means you look for what is different and what is the same.”*

T: *“When you compare what are you looking for?”*

S: *“Different.”*

T: *“Think again.”*

S: *“Same.”*

T: *“How do you know if you are looking for different or the same? If you compare you are looking for things that are the same. And contrast looks for things that are different.”*

Suddenly students in a nearby class break out and sing Rudolph the red nosed reindeer. A student comes in and reminds her that all the students in this mini-school are singing songs together in the hallway.

Excerpt two

In another fourth grade classroom, students are preparing to practice a listening passage. The teacher explains how the test will be given, what to expect and what to do to do well. She tells them: *“The teachers who score the tests get these big boxes and they sit*

there and score test after test after test. So you want your paper to stand out. Make it interesting. Don't make the scoring teacher fall asleep."

A student asks: *"I don't get it. They say take the time but then you don't have a lot of time to do it right."*

T: *"I will tell you how much time you have left."*

(...)

T: *"What is the listening part about?"*

S: *"1st time you listen. 2nd time you write notes."*

T: *"What's a good way to take notes?"*

S: *"Short bullets."*

T: *"What's a good way to set up notes?"*

S: *"Beginning, middle, and end."*

T: *"What's another way?"*

S: *"Problem and solution."*

T: *"What else?"*

S: *"Setting and plot."*

T: *"How are you going to decide whether you are going to use a problem and solution or a beginning, middle and end?"*

S: *"From the story."*

The teacher has them read the directions to themselves. Then she reads the story once. She reads it again. The students take notes, some consistently, others occasionally. One boy doesn't take any notes. He puts his pencil down and closes his booklet. He just sits there. The teacher ignores him.

T: *"Remember how important your notes are because they and your memory are all you have to answer the questions." Ok, let's do number 30. I won't be able to read it next week. You'll have to do it yourself."*

The question asks them to think of a new title for the story and then they need to explain why their choice fits the story and be able to back up their choice with details from the story. The story is called Spaghetti.

She talks to a student:

T: *"Should you start your reasons with because?"*

S: *"No."*

T: *"Now explain why that's a better title. Can you go back and look at your notes?"*

S: *"Yes."*

T: *"Can I read the story again?"*

S: *"No."*

A boy hasn't yet written anything. The teacher asks him what would be a good title. He shrugs his shoulders. She tells him not to think too much because he'll run out of time. She explains to another student that her answer only states why the other title

wasn't a good title but that's not the question. She reminds her to tell why her title is a good title and tell why. Reading over another student's response, the teacher asks her where she got her idea. The girl says from the picture [which is of a cat]. The teacher says it has to come from the story not the picture. Another boy says he got his idea from his imagination. Again he is told he needs to get it from the story. The students are getting pretty restless. As the teacher goes around she repeats this point, "*Did you use something from the story?*" Even after she says this out loud, she seems to need to say it to the next student and the next. Several students are confused or drawn to the picture of the cat. Finally she directs the whole class: "*Boys and girls, I've been able to go around and look at everyone's. The problem is next week I won't be able to go around and say anything. I also noticed that a lot of people used the picture. If the directions tell you to look at the picture then you look at it. Otherwise you use the story. This picture is just a decoration on the page. This picture has nothing to do with the story.*"

The teacher has a student read her answer to the class. She wants the rest of the students to listen for two things:

1. Did she provide a title
2. Did she use the story to explain her title

The girl reads: "*I think Lucky because Gabriel is so lucky because...and because.... So I think this story should be called Lucky.*"

T: "*She used two examples to explain the title, two reasons from the story. Would anyone argue with that title?*"

No one does.

Excerpt three

Teachers also prepare the students by giving them strategies to help them self-assess. For example as part of an activity around reading the book The Velveteen Rabbit, a fourth grade teacher passes out a worksheet and tells the students that she is going to show them how to give a response that is a 4, or a 3, or a 2. She directs them to put a 4 on the back of their worksheet and an arrow next to it.

T: "*If I'm going to write an answer that is going to score a 4, what does it need?*"

S: "*Answer complete.*"

S: "*Neat.*"

T: "*I agree but I wouldn't worry about neatness first.*"

S: "*Topic sentence.*"

T: "*Yes you need to have some sort of topic sentence. You need to remember to restate the question. What else?*"

S: "*Details.*"

T: "*Yes, details, details, details. Where do you get the details?*"

S: "*In the book.*"

T: "*Ok, it's complete and it has a topic sentence. What else will people scoring be looking for?*"

She reminds them about the DLA (Daily Language Activity) hints she gives in the morning. These are punctuation, spelling, capital letters, and correct grammar.

(...)

T: *"Leave a space and put a 3. What is going to make a difference between a 4 and a 3?"*

S: *"One of those things is not included."*

T: *"Everything needs to be there but will it be mostly complete. Will it be perfect?"*

S: *"No."*

T: *"Will your spelling be perfect?"*

S: *"No."*

T: *"Ok, what might a 2 be?"*

(...)

T: *"So what you have to start doing in your head is thinking how much information is in the story and how much you include in your answer."*

They create a scoring rubric:

4 > complete, topic sentence, details, sp/p/c, correct grammar, ending sentence

3 > mostly complete, topic sentence, details, sp/c/p/ grammar, ending sentence

2 > partly incomplete

1 > little information

(...)

The teacher then goes over some examples with them:

T: *"Now look at your question. I'm going to write my answer and I'm going to ask you what you'd give it."*

She writes: He feels plain. The students give it a one because it is too short. We don't know who he is. The teacher comments on the need for more details.

She writes: The Velveteen Rabbit feels plain and ordinary. The students give it a 3. The teacher disagrees and gives it a 2. She says she's missing details from the story. *"Have you proven it from the story?"*

She writes: The Velveteen Rabbit feels plain and ordinary because all of the toys make fun of him. The students give it a 4 but the teacher disagrees and gives it a 3. *"Did I give you specific words from the story to provide evidence?"*

She writes: The Velveteen Rabbit feels plain and ordinary because all of the toys make fun of him. For example, the expensive toys snub him and make him feel commonplace. She tells them this response would get a 4.

One girl while copying the answer says she doesn't agree that all the toys make fun of him because one doesn't. The teacher agrees with her and changes the word all to most on the board.

In the end however, test preparation is just that, preparation. What happens on the day of the test lies in the hands of children. As this teacher states: *“Today after doing a little practice with them I finally just looked at myself and attributed it to the day off yesterday and a little bit of burn out on the kids hopefully. But I can’t take the test for them and I can’t make them go back and check their work any more than they’re doing right now. And I turned to my para and I just said, I can’t make them go back and proofread anymore. You can sit there and say proofread, make sure your sentences start with an uppercase letter and end with a period. Well they sit there and stare at the paper and they think they’re doing it, but they’re not. And you can’t go over there and say, ‘that needs a capital letter. That needs a period.’ And you’re sitting there and you’re saying what more can I do?”*

The pre-testing years

We had a meeting the other day about eighth grade tests. Because I’m teaching in sixth grade I don’t have any tests that I have to prepare them for this year. But you’re always aiming for something up at the higher level. So regardless of the fact you’re in third grade, you don’t have the ELA, you’re still having to worry that next year they’re not going to be prepared for it. So you’re always preparing. It doesn’t matter what grade.

(Willow teacher)

Under the pressure of maximizing the performance of fourth grade students the foci of the early years has changed as well. Teachers have commented that it is not so much that they’ve changed their curriculum. But they have changed what they prioritize in their curriculum and how they present this information to students. Second and third grade teachers make conscious decisions to focus their teaching in ways that will benefit their students on the ELA and math tests in fourth grades, and 6th grade teachers do the same for 8th grade. A teacher comments: *“At the faculty meetings, the principal is good about emphasizing that it is not a 4th grade test. It really starts back in kindergarten. And so he’s very good at communicating that to the faculty, that it’s really everybody. And we have the Early Literacy Profiles this year and within the past couple of years they adopted the reading series to get everybody on the same page. So I think they’ve tried to adopt curriculum and do it more building-wide to get everybody on the same page.”*

A third grade teacher talks about how she has changed the way she teaches as a way to better prepare her students for 4th grade: *“Especially with reading, I do a lot of note taking, a lot of graphic organizers, and I don’t think if it wasn’t for the test that I would use them in such detail. Sometimes they are not necessarily ready for it when we do it and it’s very guided but just that exposure to it. I do much more in-depth note taking than I would typically do with third graders because they have to be able to do it when they get to fourth. In math, and I get stuck between a rock and a hard place every single time. I feel like I’m rushing through my math series, that there’s times that I know they don’t necessarily get it because there’s so many other things I have to get through. And*

I've gone to the fourth grade teachers and said 'help me out. Do you want them to at least be like oh I've seen that before and not necessarily have it down or do you want them to come to you with less and have it down?' And they've said that they would like them to have at least the exposure. So I feel like a lot of times I'm not giving them the time to master a concept in math before I'm moving on."

During a reading lesson a second grade teacher includes, among many other activities such as vocabulary and comprehension, an ELA type question and answer session creating a story map with her students. They have already listed the title, author, and illustrator. A fire drill interrupts the process. When we come back inside the teacher tells her students that she is extremely proud of them. They did what they were told, they waited patiently and they were quiet.

T: *Ok back to Bully Trouble. There are several characters in the story. Who are the main characters?*

They talk about several of the characters distinguishing between main characters and important characters, adding these to the web that the teacher is creating on the blackboard (see Figure 2 below).

T: *So we have the characters. Where does the story take place? What is the official second grade term for that?*

S: *Setting.*

T: *Excellent. So where is the setting?*

S: *Outside.*

Waits expecting more. Another student answers.

S: *In Robby and Arlo's neighborhood.*

T: *Very good.*

They list and talk about the characteristics of this neighborhood.

T: *As with most stories, there is always a problem.*

S: *Big Eddie was taking their soda.*

S: *The problem was there was a bully teasing them.*

The teacher and students talk about the problem at length sharing different characteristics of the relationship between the bully and the two main characters and then create a succinct statement as the problem.

T: *Good choice of words. Big Eddie is a bully to Robby and Arlo. He teases them. He pushes them around. Is that the problem?*

S (several of them): *Yes.*

T: *Solution. What's a solution? The solution is how they solve that problem.*

S: *They made a bully special with vinegar (several students contribute to the ingredients), and steak sauce, and prune juice, and chili hot sauce and put it in a soda can.*

STORY WEB

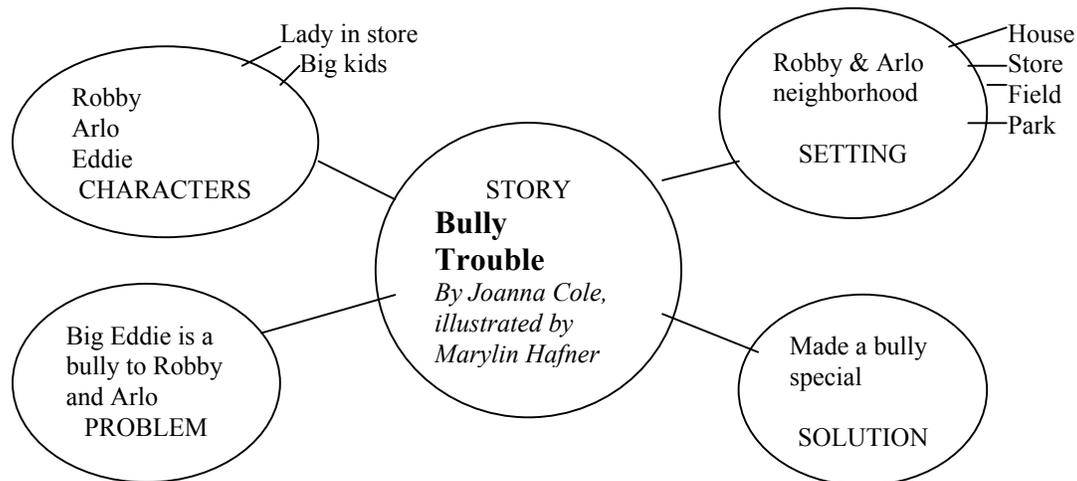


Figure 2: 2nd grade story web for **Bully Trouble**

The second grade teacher shares how some of her goals go beyond the objectives for the curriculum in second grade: *“Always in the back of your mind what they will be doing next year, the independence that they are going to be doing in third grade. You know where they need to be and you’re always leaning towards that. Now with the 4th grade ELA test, in order for them to map out their thoughts and their ideas and be able to write within a sequence of, and be able to recall details from a story and plug it into where they want it. And if I can give them any kind of small strategy that they might be able to pull and say “oh I can do this,” or “I’ve done that before.” Any strategy at all that will get them to be able to comprehend and be able to write on and expand on. I have a list of comprehension questions that we will work on tomorrow and one of them is the multiple meaning of the word ‘can.’ Did you see them pull that out? The word ‘can’ can have multiple meanings. If I can give them the strategy to be able to use the context clues to understand which word they’re reading then it’s just another strategy for them. Have I changed my strategies in teaching because of the ELA? Not on purpose but yes I think I have. You know I’m including more writing. I’m including more story webs. I’m including editing and all those things that used to be considered writing but are now all part of language arts and you’re including them within, you know the language arts block of time.”*

Supporting change

All eyes on ... fourth grade!

Despite the efforts made by K-3 teachers and the rhetoric of a shared responsibility across grade levels, in practice the teachers generally feel that most of the energy and focus from administration is on 4th grade. A teacher shares her perception of these practices: *“They are more concerned about third and fourth grades and seventh and eighth grades. The push was to lower the numbers in kindergarten, first, second*

grade, in the early primary years. Now the push is to lower the numbers as they get closer to third and fourth and seventh and eighth because that's where the testing is. Also in the past we looked at the primary grades as being the formative years. And now we're looking at the focus surrounding the test years rather than the focus being in years when basic learning is formative. Not that we ignore K, 1, and 2, but the focus is not placed in those areas. And often it's at the expense of other programs when something is created, we borrow from Peter to pay Paul rather than looking for funding sources elsewhere."

Part of the challenge of 4th grade is that teachers need to both teach several new units well enough for the students to be able to answer questions on the test, and review everything they have learned up until then. The tests cover all the years prior to fourth grade, and it doesn't really matter if it was taught well in 2nd or 3rd grade because retention of all the vocabulary or concepts isn't likely to occur. But for fourth grade teachers it is not only the test and the curriculum that is challenging to juggle, but the scrutiny they experience from administrators, the community and the media. As teachers try to make sense of their experiences, they discuss how the administrative attention has both good and bad effects:

"I honestly feel that the test is a realistic test. I do not think that the demands are outrageous. I do think it's a difficult test for the majority of our population. And I think that if we start making changes at the lower level eventually it won't be as difficult as it is right now. But around this time of year, starting in November, it's that mad dash to 'what are we going to do to get these kids ready?' Sometimes I just feel like saying, stop! It's not what are we going to do to get these kids ready. What have we been doing and how can we keep doing that? And when we sit in the fourth grade team meeting room and we're all sitting there and hear from administration, 'What are you going to do? What can we do?' And it's like saying it's not completely our responsibility. Why aren't the 3rd grade teachers sitting here? Why aren't the second grade teachers sitting here? And I'm being very candid, why aren't the 1st and kindergarten teachers sitting here? Why aren't the parents sitting here? It's not what can we as 4th grade teachers do? And that's what happens every month we have a meeting and we hear the same thing."

"And I get a sick feeling in my stomach because it's like how am I going to get 4s out of some of these kids."

"Right and then I find myself justifying well, this child can't possibly get a 4 because. And I will work really hard to get this child from a 1 to a 2. And then I say to myself, why am I justifying this? This is this child's ability level and the test really should be a true reflection of what that child does. And considering the results of our test this year, it's been a real eye-opener in terms of where our kids have come from and how far they're going to go and they've made tremendous progress, but it's not 3s and 4s. Those who were non-readers will now be 1s and 2s and I feel great that that's where they're going to be. But they're not going to be 3s and 4s and that's the reality of it this year."

“And it’s not a tense meeting in and of itself. It’s more of a what are we gonna do? And we are doing what good 4th grade teachers are supposed to do. We’re teaching the students the curriculum. You can’t ask us to make up for the fact that this child is deficient in this skill and has been since kindergarten and there are just certain things that are beyond the 4th grade classroom teacher’s control. For example, a child whose testing request has been in for 3 years but the parents refuse to sign the paper and I’m asked what am I going to do? Well first of all I can’t contact the parent because there isn’t any phone. Mom doesn’t respond to any letters that I send home and I asked guidance to intervene. I can’t do anything more.”

“And I think that’s what she means when she says you walk out of there feeling sick. You’ve done everything you can in terms of the steps to follow in order to make sure this child is getting everything they need and there are kids who will be listed on a student-at-risk form for years and years and years and nothing happens and boom! They come into fourth grade and all of a sudden they’re filling out testing requests for them and now they’re giving them remedial services, all because somebody looked at their paper and said ‘Ooh, they may not pass the ELA.’”

“Being in fourth grade, it’s almost to an advantage too. If I need materials and if I say, ‘oh, it’s test related,’ then I can get them. If I have a child that I need to have looked at, ‘oh, it’s fourth grade,’ seems to be there’s more of an emphasis on something, whether that’s good or bad. You know, the school’s ready to give you materials if it’s test-prep related.”

Focusing on 4th grade is one way that a school might address the testing requirements placed on them and is not an unusual response. Teachers understand this, but at the same time, would prefer a less central role in the accountability system. They know, just as administration knows, that the results on the 4th grade tests are impacted by much more than what occurs in the 4th grade classroom.

Meeting the needs of students with special needs

It’s hard to prepare the special education kids for the assessments. They need consistency and structure. As soon as the tiniest, tiniest thing changes, they’re very needy in that sense. That so terrifies me about them going into fourth grade because their independence even on a very simple task, read the directions and complete it, on something that’s not even a difficult academic task, just the fact that they have to read the directions and use them. I have kids that if they’re coloring something and it says ‘color in blue’ and they can only find a blue-violet crayon, have no clue what they should do. I mean they just don’t have that. If they only could find a blue-violet they would sit there with a blank page

because they don't have blue. They can't just realize that it is blue. That they cannot do.

(Willow teacher)

Students who are considered for special education or other services are referred to the Child Study Team (CST), a group made up of guidance counselors, remedial reading teachers, and teachers, who then evaluates the student and gets parental permission for testing. The purpose of the Child Study Team is to assess the academic and psychosocial needs and strengths of each child. It takes into account the multiple issues affecting students who are or may be considered at-risk. This includes absenteeism, truancy, psychosocial and emotional needs, learning disabilities, as well as a range of behavioral issues. That process may lead to a child being referred to the Committee on Special Education (CSE), or recommended for remedial reading or math services, speech, counseling, English as a Second Language, and AIS (Academic Intervention Services), or it may lead to retention as being one possible intervention used to help a child succeed in school.

Like many city schools, Willow Valley Elementary has its share of discipline issues. A school-wide discipline system based on Lee Canter's Assertive Discipline provides the basis for dealing with discipline issues. Teachers are also required to create a behavior management plan, either as an individual or by mini-school. This does not always mean, however, that everyone agrees with how discipline is being carried out. For example, several teachers state feeling challenged with the behavior issues they encounter at Willow Valley. Teachers comment that their concern is not that the administrators are not following the school-wide discipline plan, but that they do not understand how disruptive an out-of-control child can be. What happens, teachers explain is that the child is often sent back to the classroom when he or she is still unwilling to cooperate. Teachers feel that they often need to rely on each other to provide time-out space for a disruptive child.

Discipline issues affect instruction in many ways, as this teacher comments: *"I find that a lot of my time is spent on discipline. And it's like teaching these kids manners and respect, and how to be good kids and get along with each other, you know, just respect each other. So much time is spent doing that. And I try to implement it through the lessons I teach, whether it's the DLA [Daily Language Activity] or in character education. It's like all of these things that I've got to do and there's no time. You look at the clock and the day is gone."*

And parents agree that discipline is a sensitive and important issue that needs to be addressed. Many commend the teachers for doing an excellent job in a difficult context. *"This year the teachers that we have are absolutely fabulous. They have large classrooms and the 6th grade class is very difficult. There are difficult children in it. But the teacher is, she's very, very good. She's very in touch with me."*

While others are concerned that their child's education is suffering because of the amount of attention that gets spent on behavior issues. *"My son can benefit from a more one on one type of atmosphere. But I can tell you right now the amount of time the teacher spends regrouping kids, dealing with discipline problems, and all that other stuff, they don't spend nearly as much time. You can get done in two hours at home what they can get done in a whole day at school."*

And while students with behavior issues are not necessarily or even usually students who receive special education or remedial services or vice versa, teachers are not always satisfied that the needs of the student population whether behavioral or remedial are met effectively within the school. They disagree however, as to where they feel more resources should be directed. Again, some teachers would like to see increased special education and remedial resources, while others would like to see a stronger emphasis placed on strengthening integrated and intact classroom environments. One teacher comments: *"One thing we do a lot in this school is identify children. Student behaviors are often textbook indicators: high in math, low in reading, attention problems. Or is the answer that he is showing that he likes to do what he's good at and not wanting to do what he's not good at? Identifications may not be the answer. His effort correlates with his grade and effort in reading correlates with reading grade. Is this human nature or really special education?"*

Remedial services and special education teachers get caught in the middle of the assessment-driven reform efforts. On the one hand, they know their students need to perform well on these tests and they wish to assist classroom teachers in making this happen. On the other hand, as teachers specialized in providing particular resources, they're job is to focus on the needs of their students in those particular areas. One side effect of needing all students to take the exam is that the focus of these services, the need to focus on reading and learning strategies can get diverted. One remedial teacher describes how this might happen: *"The first year the state tests came out, maybe the first two or three years we taught to the test. And then in that time we got our new reading series and it really goes along with what New York State is asking so lately I've been phasing away from teaching to the test and just really focusing on remedial reading. When we taught to the test we used the sample tests they provided and just used any tool that was out there that copied what the test was like. So that's how we taught to the test. But I don't really focus on that anymore. I do not like teaching to the test. I really do reading. The first year it was just natural, it was the first time we had seen it. There really was no training, so we were unsure, so it seemed natural to kind of work towards it. It went well the first year so we stuck with it the second year and then it seemed like we shouldn't be doing it. I mean the focus is that my children can't read, whether they know what the test looks like, I still need to get them to be able to read it. It gave them the familiarity with the test and they were comfortable in the setting and they knew what was expected of them, however, they still struggled with the reading. So in the last couple of years I kind of phased away from teaching to the test and really focus on their reading and ability to understand."*

On top of placing additional requirements for students, the state's accountability measures have added new challenges for teachers working with special needs' students. A special education teacher comments on the difficulty of the tests for special education students: *"On all three tests, the ELA, the math and the science, there are parts of it that are well beyond what life will be, well beyond what these kids are capable of. But I think it's part of educating them for the real world. They're going to meet challenges that are beyond what they can do. We don't want to dummy down the test to the point where it meets their level because life can't be dummy downed to meet their level. And in many cases the kids will surprise us and rise above where we expect. But the frustration that they meet along the way is the sad part. And the pressure that they feel in handling something that is beyond them. That's why the expectations need to be raised right from the beginning, right from pre-K on up. And that is easier done when you have a very rich environment at home, and many of our kids don't have a rich environment, they don't have an average environment, much less a rich environment. So I think the difference between the performance levels in (wealthier districts), there should be some provision for home environment. We look at it for the non-identified kids but we really don't look at it for special education students. Special ed is considered special ed. If you are in a district with high free lunch rate, you know your poverty rate is high. You look at your total population, and you make different judgments. But when it comes to the special ed population because they're handicapped, they're looked at unilaterally and that's a very difficult situation. Once they're considered special education they are considered to be more similar than dissimilar, those economic factors aren't seen as having as much impact."*

State guidelines for modification add to these issues by being so vague that teachers may end up interpreting them differently. A teacher comments: *"A good example of this problem is that many schools up until this year have actually read the reading comprehension part of the ELA to their special education students. They interpreted 'test may be read to students' in their IEP as meaning all tests. You totally invalidate a reading comprehension test by reading it to the child. But in the worry, in the fear of poor scores and what that would entail, they over-interpret the modifications of reading tests to include reading comprehension. And that no longer can happen, the mandate came down from the state that they can not do that."*

She goes on to explain that thankfully Willow has not been one of the schools that has been reading the comprehension part of the ELA. If they were, she explains, it would probably cause a significant drop in test scores. As it is, they should remain fairly steady. But even with clear guidelines, the pressure to help students do well on the assessments may push teachers to stretch their interpretations beyond what is acceptable, as one teacher comments: *"There's all kinds of subtle ways that consciously or subconsciously we can give hints to students and I'm very aware of not doing that. And some people, not necessarily in this district, use the excuse that the test is well beyond the students' means so we need to give them everything that they can get. And they extend that beyond what is ethical and legal."*

These issues may be one consequence of having the state mandate mandatory levels of achievement for all students and pushes shared responsibility to mean much more than providing a quality education to all students. This is because input into deciding what counts as a quality education lies outside the reach of most teachers and administrators and distorts other ‘measurable qualities’ occurring in schools.

Furthermore, districts are not provided with any additional resources to implement the many assessment requirements placed by the state. The principal talks about this issue: *“I don’t think with the different things that have come down from the state that they’ve provided all the resources that they can possibly provide to help us achieve success. You know there are certainly a lot of un-funded mandates that the state does all the time, not necessarily in testing but in a lot of different areas. But I think they have effected the benchmark some place and they decided now they’re asking districts to get there and the ones that can’t get there unfortunately have to fail before they get some extra resources from the state.”*

It should be noted that teachers are not concerned about the quality of the special education and remedial staff but about the impact an overburdened system has on their ability to carry out that quality. And teachers are not critical of the district policy towards integration feeling that overall it has benefited both regular and special education students. Teachers have, however, expressed concern that the new state requirements have had the effect of placing additional demands on an already overburdened system in ways that could backfire on the capacity of the total system. For example, there is a general sense that because of new state requirements more students are being referred to the CST committee for help. Without forming conclusions about the motive or actual practices in place, teachers have expressed concern that recommendations for special education, remedial services, and/or retention have increased as a result of the pressure to perform well on the state tests:

“Unfortunately, sometimes the goal is to have students classified prior to taking the test so that modifications can be made. So that the child who probably would have scored a 1, with the modification could score a 2 or maybe even a 3. If the child needs the services that’s wonderful, get the child the services, that’s appropriate...”

“But get them when they need them in the beginning...”

“Right, and not put a Band-Aid on it because there’s a test coming up. Because that’s not fair to the child, because all of a sudden you have modifications that he or she hasn’t had a chance to adjust to or practice.”

What is frustrating for teachers is not that there is a lack of response to the needs of students or that the push to get students identified is uncalled for. The concern is that there does not seem to be a coherent, well thought out plan for addressing the needs of students at risk. They feel that the district seems to respond and react to individual needs

as they arise but fail to plan for the fact that these needs exist in a variety of ways across the entire population of students.

Teachers, administrators and parents agree that the percentage of students who qualify for special education services is high. Table 2 provides the number of new referrals to special education over the past five years. The Director of Programs and Services explains: *“The state always looks at 11% as a guideline for districts. So if you go over 11% what they do is they want to look at programs and see what you can come up with and how you can keep the classifications down. Our district right now is running between 16 and 17%. We have approximately 300 special education students identified.”*

He explains that a reason for these increases is probably related to concerns that students with learning disabilities will not be able to score well on the state assessments. Helping these students feel successful would be a primary goal: *“I think part of the reason is the concern with the state assessments. So I think there is a tendency to at least go through the referral process. Even the 504s we’ve seen an increase in. And the reason for that is so they can provide some testing modifications for the students. That’s really what the benefit is. That it will help the students during the state assessments.”*

Table 2: Number* of new referrals to special education

Grade	1997-98	1998-99	1999-00	2000-01	2001-02
K	5	4	13	15	2
1	5	9	5	9	2
2	10	16	2	9	6
3	5	11	10	10	9
4	7	14	7	5	4
5	5	4	3	15	3
6	1	7	2	10	4
Total number of students receiving special education services:					
In-district K-6	123	124	142	146	-
In-district 7-12	75	65	90	96	-
District Total	198	189	232	242	-

*Note: These numbers are misleading since referrals are counted from September to September when they get processed and staffing determines that, while the numbers of students receiving services are counted on a December to December calendar year.

The pressure to have to perform well on state tests pushes systems to employ strategies that will hopefully result in higher test scores. As Willow Valley has shown, systems can use a multitude of such strategies. What also seems to happen is that systems are most likely to increase the use of strategies that are already in place rather than create new ones. Relying on remedial services and special education is one such strategy. Another strategy is retention.

Willow Valley has a well-developed retention process based on the School District's Promotion and Retention Policy adopted in October 1998. Starting in October of each school year, teachers submit a clear documentation and referral trail to the Child Study Team so that in May when a decision is made, the justification is clear. CST uses as a guide Dr. Wayne Light, Light's Retention Scale (LRS), 1991 Edition, Academic Therapy Publications, Novato, California. The Light's scale provides a comprehensive guide and is based on the positive and negative effects of retention found in the research. In addition, the committee weighs the strengths and weaknesses of each situation using such criteria as the child's size, maturity and immaturity, their grades, their home life, and their ability to perform in the next grade level, especially if going into fourth grade.

Most teachers have not expressed being against retention, as this teacher comments: *"It's been a wake-up call to end social promotion. We are not doing the kids a favor promoting them so they fail."* However many teachers are still confused about the process they are required to follow and are concerned that due to the pressure of the state tests, the number of retentions in second and third grade have increased. Before September 2001, 89 at risk students were discussed and 51 retained; 42 of these were K through 3rd graders. Several teachers share their perceptions on this issue:

"If the state has decided that 4th graders are ready for serious assessments, then it is no wonder that schools are going to reevaluate each student before that grade. Is this a negative reaction to state pressure that will hurt kids by being identified unnecessarily or is this a wake-up call where before kids were being socially promoted and failing later on? Looking at their needs earlier on might be a good thing."

"One of the things that I'm most uncomfortable with is the effect the testing has had at the mid-year, when it's time for us to decide is this child going to meet the criteria to move on to the next grade. And that criteria has drastically changed in my opinion. It is no longer weighted very heavily, when you as a professional say, 'I know the solution for this child is not retention.' Especially on the third grade level, the message is loud and clear."

"I have personal concerns about our programming, that one way to achieve the standards is not getting kids to the 4th grade who are not ready to achieve."

"It's probably different for a fourth grader."

"They don't want to retain them. They want you to retain so they're not going into the fourth grade and when we start making those decisions based on, and ask the Child Study Team 'how would we classify a child?' 'Well, what about the test?' 'I don't care about the test. How do we classify? Don't think about the test!'"

"The test permeates everything."

Early literacy assessments

One new development for supporting change that teachers have mixed feelings about is the implementation of a new series of Early Literacy Assessments for students K – 3. New York State requires that schools develop some form of literacy assessment for the early years as one way to have a better grasp of what students are learning before the 4th grade. This policy is supported by the passage of President Bush’s No Child Left Behind Plan, which requires assessments in reading and math K through grade 8. So it is no surprise that the Willow Valley Literacy Committee chose to develop a series of Early Literacy Assessments for grades K through 3rd grade as a way to better assess the early literacy skills of Willow Valley students. These assessments are considered diagnostic and test students’ literacy skills based on the reading series that the district has adopted and uses the format of the ELA. Table 3 provides a history of the district and state tests given to Willow Valley students.

The curriculum coordinator explains: *“These assessments will be administered in the spring of each year and will document student progress over time. We believe these tests will help identify those students who need Academic Intervention Services in order to be successful in meeting the NYS learning standards in ELA. They will help to prepare students to achieve proficient and mastery levels on the NYS ELA assessment at grade 4 and grade 8.”*

Table 3: History of district and state assessments given to Willow Valley students by grade level

	1997-98	1998-99	1999-00	2000-01	2001-02
K					Early Literacy Assessment
1	Stanford		Terra Nova	Terra Nova	Terra Nova Early Literacy Assessment
2	Stanford		Terra Nova	Terra Nova	Terra Nova Early Literacy Assessment
3	Stanford		Terra Nova	Terra Nova	Terra Nova Early Literacy Assessment
4	Science	ELA Math Science	ELA Math Science	ELA Math Science	ELA Math Science
5	Stanford		Terra Nova	Terra Nova	Social Studies Terra Nova
6	Stanford		Terra Nova	Terra Nova	Terra Nova

For the past two years the Terra Nova was supposed to be fulfilling this function. But bringing in outside tests like the Terra Nova are costly and may or may not be good overall measures of the local curriculum. The principal provides his thoughts on how the district’s literacy assessments are supposed to change that: *“The Terra Nova as a standardized test is not taking any specific curriculum. I think our literacy assessments*

are taking all language arts reading curriculum and major backbone parts, structures, and making sure that everyone pretty much gets in step with that. Teacher X is never going to be identical to teacher Y in what they give the kids. But we want some things that no matter where they go in this school they're going to have the same material covered and hopefully the same level of mastery that will be a pretty good predictor of their future success as they go up the ladder. What I want teachers to identify is that kids coming from you in first grade to me in second grade have to know this body of knowledge or they're not going to be able to make it here in second grade so this is what the minimum is that they need to be successful. Those are the things that we have to define and this is the process that I'm hoping that they can focus on. And also I don't want to have to wait until the 4th grade test to know that some kids are really at-risk. So that's why we put them in."

The district's curriculum coordinator adds that by developing their assessments using pieces from McGraw-Hill publications, the company that's actually writing the state tests, that it should be well aligned for the purpose of predicting how well their students will do on the state tests. But she states it was also guided by, *"discussions we were having in the district on committees, looking at the math 4 test and saying you know, these kids didn't do well and it wasn't because of math, it was because of literacy issues."*

What happened, however, is that the teachers were presented with the literacy assessments during the March superintendent's conference and told that they would have to administer and score these before the end of the year so that the results could be added to the diagnostic tools to determine student retention and need. The teachers' response was not exactly receptive. One teacher explains: *"This year in March we were handed a Literacy Assessment and now every K through third grader has to take that and pass that before they can be promoted to the next grade. I'm sure the literacy committee worked very long and hard on this. From a reading specialist's standpoint it's a fabulous measurement for the reading, the writing, the listening. It's all there. From a classroom teacher- 27 children- perspective, you just gave me a whole other set of work for the last six weeks of school that I didn't need. And I'm not even sure how you are and why you're going to use it. Since April I have tested my students in Terra Nova's, Theme 5 reading test, Theme 6 reading test, Literacy Assessment test, not to mention you're other benchmark tests in science, social studies and math. The last 8 weeks of school, my children have done nothing but take tests. They've done nothing but take tests. Fabulous if you're going to use these assessment tools to see where my program is going and where I should be. But if no one is going to use these assessments for where we're going with these kids?"*

Teachers are concerned that Willow Valley's retention process begins in October and cannot be altered based on a single score in late May. So the lack of clear guidelines on how to use the results of these assessments for promotion decisions, along with the

lack of preparation for implementing and scoring them, is just one more last-minute decision that several teachers feel was not thought out properly.

What is frustrating for some teachers is not the additional assessment piece, but the lack of an overall plan for how these assessments will contribute to the curriculum and pedagogy offered students at Willow Valley. And there is no reason for them to believe that any comprehensive plan is in place since other assessment tools that are given to students have not been followed up with any analysis either. So the teachers ask, Why add a redundant assessment? Why not look closely at what students are doing on the benchmark tests in each subject matter and build programs based on that analysis? A couple of teachers comment:

“The literacy assessments were driven by the idea of assessment driving the curriculum. I don’t think it’s that much different from other stuff that we are doing. So it is really one more redundant assessment. The advantage is that we are all using it. It’s unclear to me and it’s unclear to others, what are these for? Assessment is only valuable if you are going to use it somehow.”

“I don’t see where administration is helping us take the results of our tests, clearly identify what our needs are, and facilitating moving on, meeting those needs. The Willow Valley community will always have a chunk of people where education is not their top priority. There’s programming that can address that. I don’t see us really brainstorming to do that.”

The principal is cognizant of the need for more data analysis initiatives across the district. What makes this difficult he explains is that in a small district, most staff members are already part of several committees and they simply don’t have the people power it takes to do what needs to be done. The work the district is doing with aligning the science curriculum and the literacy assessments are just beginnings, he states. Other efforts will follow as the time and resources are made available.

Going the pendulum

The reasons for creating alternative programs and services for students is to improve the academic performance of all students and especially those considered at-risk. School boards, state departments, teachers and administrators share these goals. Holding measures such as test scores and improvement benchmarks such as the Adequate Yearly Progress (AYP) set by the state creates responses that can have conflicting effects even when the actual goal, such as improved test scores, occurs.

The previous sections have described numerous decisions made at the district and classroom levels that have impacted the education students receive at Willow Valley Elementary. Textbooks have been adopted, professional development choices decided, and new instructional practices delivered. Teachers comment that like other districts they

too have gone the pendulum from whole language to basals and from manipulative math to textbooks. The efforts made at the district and school level are similar to other districts facing the same state requirements. And like other districts, personal and professional tensions that existed before the advent of state testing continue regardless of other changes occurring. But Willow Valley also presents its own unique features, which may play a role in how teachers and administrators have responded to these changes.

Willow Valley has a history of encouraging a student-centered, integrated, and individualized instructional approach for students. Students who may not succeed as well as others or in the same way are provided a variety of classroom situations to choose from. Success was determined not by some outside measure but through the teachers or the Child Study Team's determination of the strengths and weaknesses of each individual case. State testing has affected this practice in more ways than one.

The first evidence with how this practice has been affected is with the language teachers and administrators use to talk about their students. While teachers and parents want to maintain an individualized approach, the testing pushes the entire system into tracking students along one criteria: their potential test score. No longer are students discussed as boisterous, creative, needy, arrogant, quiet, analytic, or literal; instead they are either 1s, potential 2s, 2s, potential 3s, 3s, potential 4s, or 4s.

The effect this kind of labeling has on how teachers measure-up to administrative evaluations is clear. Teachers, especially third and fourth grade teachers, have commented that they are assailed with messages about where to direct their energies. Many of them state feeling pressured to identify potential fours in their classrooms and work directly with them as well as identify failing students to get assistance or services that might potentially raise the outcome for that student.

"It hurt our relationship when the administration made it clear, you know we want to get those twos to be threes. We want those threes to be fours. We don't want any ones, ok? We have those same goals but at the same time I have several remedial reading students in my class and there's only one remedial reading teacher and she can only take them three days a week. And two days next. That has not changed since the standards have increased. Nothing has been done to meet the needs of those children. And all I hear from administration is you meet the scores, the scores, the scores. But I don't see anything done for the needs of my children. We have a heavy special education population in our district, and while we appreciate administration's hands are tied, the pot's just so big, we still have those stresses on us, and we don't see the standards addressing programming needs and staffing needs and things like that. We haven't seen it driving that."

"We've seen it driving us personally."

“But we haven’t seen it involve what’s offered to our kids here. It’s still us who are providing everything.”

The unintended consequence of this kind of talk is that it seems to limit the way teachers and administrators talk about the needs of their students and therefore takes away from other important and essential conversations about the teaching resources in meeting the educational needs of students.

Another effect testing has had is that it has made the process of an integrated, differentiated instruction not only more difficult to accomplish but somewhat counter-intuitive. What is the point of using alternative materials or approaches with children who are not succeeding with the regular curriculum, if their successes get reduced to performances on state developed tests, tests that may or may not be valid forms of measurement for those particular successes? So while instructional differentiation is encouraged in theory, the state assessment system works against developing a variety of approaches in practice. And teachers get frustrated when attempts by administration to help only complicate the requirements they are trying to negotiate, rather than clarify them. As this teacher explains:

“I think there’s an attempt by administration to provide us with resources but, it’s not what we need. You get a teacher aide to help you with your needy students for a half an hour. By the time you touch base with the teacher aide, get your kids settled into a group and get going, they have to leave. So even if you give your aide your better students and you work with your weaker students, it’s still the same situation, they’re not the teachers and it’s really not their job to teach them the content areas. You know books will come through our mailboxes, ‘How’s this resource?’ ‘Do you think this looks good?’ But everything is an effort and nothing is a consistent start to finish; let’s try this, let’s stick with it for a couple of years and see what happens. It’s ok this teacher’s using Blast Off, this teacher’s using Measuring Up, this teacher wants to do this, and that’s fine, you have to find what works for you. But it just seems as if there’s an awful lot of, let’s try it this year, well alright, we’re done, let’s try something else.”

Unraveling the Logic of Testing

Assessment-based reform efforts present their own logic. Among other things, teachers, parents, and administrators expect them to: (1) guide and improve the content and instruction of the curriculum; (2) provide feedback on student strengths and weaknesses; (3) motivate teachers, parents, and administrators to work harder to educate children to a higher standard; and (4) provide accurate measurements of schools and districts for accountability purposes. The work done and decisions made by teachers, administrators and to a certain extent parents at Willow Valley Elementary to meet expectations and requirements both support this logic and put into question its integrity and validity.

(1) To guide and improve the content and instruction of the curriculum

Teachers and parents endorse the idea of using tests if they adequately measure the curriculum children are required to learn and know. One concern already related is that there is little evidence that test scores give teachers a useful indication of what a student knows. Another concern is the effect the testing over time could have on the curriculum, where the curriculum and the tests become one and the same thing. Parents try to determine which comes first:

“The kids are learning things that are going to prepare them for the test.”

“Are they based on the curriculum though?”

“The curriculum is based on the testing.”

“I think the main point is though that they should be educated well enough so that no matter what kind of tests they take they should do all right on it. I think within the last couple of years because of all the new standards I think this district puts a lot of pressure on the kids because they would like their ratings to go up. They would like to improve the outlook of the school and make it more attractive for people, because few will stay.”

“But that is the standard though and I can't help but think that if we didn't do that here in Willow Valley and let's not even attach our funding to it, if we didn't do that here in Willow Valley, then our kids would not stack up, the ones that are able to learn and able to make the grade so to speak, our kids are not going to stack up to the ones that are at (wealthier schools).”

The teachers at Willow Valley generally agree. They want to raise standards and scores but are not sure that teaching to the test and teaching for learning are one and the same thing, as this teacher remarks: *“I think we end up just pushing kids. Sometimes I think what are we teaching them? Are we really teaching them or are we just pushing them? And that's frustrating.”*

(2) To provide feedback on student strengths and weaknesses

Another contradiction within the logic of testing is that if the tests are benchmarks for determining where an individual student ranks at particular points in time then why are they used to penalize students and teachers? This contradiction highlights the way concepts such as benchmarks are used in the rhetoric of the state but in actuality are not intended to mean what the teachers think they mean. For the teachers benchmarks measure a set of skills which provide feedback on the strengths and weaknesses of their students. They use these to assist them in a variety of ways, for example for planning remediation or placement of individual students or for making changes on lesson plans that would affect the whole class. Benchmarks are considered points on a developmental

path not the end in itself. In other words, benchmarks help guide the overall instruction of students but do not automatically determine it. A teacher comments: *“It’s a benchmark. If a child can’t do it in fourth grade and they get it in fifth grade, well, why should we penalize them? The goal is to get them to be able to perform at this level and if it takes them an extra year to master something, that’s okay.”*

From the state department’s perspective the intention of state tests is not to provide guidance to districts, rather it is to establish a pre-determined level of performance that districts are required to get their schools to reach. The goal of the state is for all students to pass the 4th grade tests in fourth grade, 5th grade test in fifth grade, and 8th grade tests in eighth grade without flexibility or exception. Scores of 1 or 2 are unacceptable. In other words, benchmarks are the points where all students need to converge at particular points in time. State tests provide the expected heart rate. It is up to each district to determine what it needs to get everyone’s heart beating at that rate.

Teachers assess their students all the time. They are used to looking at student work diagnostically with an eye to understanding how each student thinks and learns. These assessment practices are relational and take into account multiple contextual features. The goal is to assist the students with areas they are having difficulty with and motivate them to continue to improve. They do this by accepting that people learn at different rates and in different ways. The state tests, on the other hand, do not assess the contextual features affecting each student. Rather they are simply meant to determine whether students have met or have not met a particular standard as measured by the state test. And just as the tests seem to have replaced the curriculum, parents and teachers worry that the state tests will take precedence over other assessment practices.

(3) To motivate teachers, parents, and administrators to work harder to educate children to a higher standard

As this story has shown, a concerted effort and attention placed on testing requirements can result in improving student achievement, at least the kind of achievement that is measured by state tests. The issue of teacher motivation, however, is a complex one. On the one hand, as the teachers at Willow Elementary have already stated the state standards and tests have motivated the district and school to update their instructional practices and has motivated them to replicate these approaches in their classrooms. On the other hand, having their success determined by how their students score on the state exams, means that it becomes more and more risky for teachers to try something new or find time for ‘non test-prep’ activities.

What seems to motivate teachers are the children themselves and the way they grow and develop in the course of a year. Teachers find joy in taking each child from wherever they are and moving them forward. It is not how that child compares to others that is motivating but where that child came from and where he or she is going. And as has been shown in previous sections, the state tests undermine these efforts.

(4) To provide accurate measurements of schools and districts for accountability purposes

I like doing pretty well with not having a lot of resources. If you go into the paper and look at the scores, our scores will be as high or higher than anyone in our comparison group and they'll probably have \$2000 more per kid behind each one of them and better pupil/teacher ratios too. Well if I was looking at somebody and say who's doing a better job? I would think our staff would be.

(Willow principal)

If the logic works and the tests truly measure the quality of education in a particular setting then parents and community members should be able to judge the quality of schools based on the school's test scores and the quality of learning based on their student's scores. For example, even though the Terra Nova is not a state mandated test, the goal would be similar to what this parent reports: *"My daughter is only in first grade, but last year we just moved here. She started kindergarten in a different school district, but we had done a lot of research long before she even started. I found that the Terra Nova for her school was important to us because we had moved from another district. I know that she was getting the same amount of education here as she would in another school district and she was above average in everything. So I know she's getting what she needs here and I'm very happy with this school. What I'm saying is she is not farther behind because we moved here. She's right where she left off, because it's a nation-wide test as far as I know."*

For the most part, however, parents and teachers do not feel that the test scores present an accurate picture of what the school or student is capable of and that is a serious issue when all that gets reported to the state department and to the public are test scores. One problem for teachers is that the media leaves out a proportion of the Willow population when reporting test scores. A couple of teachers comment on this issue:

"The state has these high expectations and I think that they don't take into account the home environment of the children. So many of these children come to school and they come from environments where education and learning are just not valued. And you know the day of the test comes around and I have little Johnny sitting in front of me, and little Johnny may have been beaten this morning, or he may not have had food since he ate lunch here at school yesterday, and is little Johnny really interested in getting a four on that test? Or is he really interested in coming to school to get some food and get a hug from me? So I sometimes feel like we have the people sitting at state ed in their little white offices, and not really thinking about the kinds of issues that are day-to-day issues that we deal with. And I feel that it's not really fair to them to expect these kinds of things when they're really struggling to have their basic needs met."

“And they [the media] don’t know that out of a class of 20 kids, maybe there are 14 remedial students in that class. Or, out of this class maybe 4 kids have moved in and 7 kids have moved out. I mean that information isn’t there and you don’t have these kids from September until June. One or two kids come in, one or two kids come out, and that really changes the makeup of the class.”

Furthermore, teachers have a difficult time resolving how to understand whether or not test scores provide an adequate measure of what they do. When a measurement doesn’t make sense to teachers and doesn’t provide useful or additional information about students, it is difficult to see the value in it. Several teachers discuss this issue:

“This test determines whether or not Willow Valley has done a good job.”

“Yes and no. Because if you get kids, if you get a class, you might do less work one year and the kids do better because the kids were more prepared.”

“It depends on the deck of cards you were handed.”

“And the next year you might do the same thing and not get the same results. Does that mean you didn’t do a good job? It’s a tough thing to assess.”

“Is there any other way to evaluate if you don’t use tests?”

“What if we have a kid who got a two on the ELA, but was an emotional disaster, disruptive, but during the course of the year in their behavior, in their courtesy and respect improved tremendously. Are you not a success then? Did that kid not improve? Are they measuring that?”

“And how do you measure that?”

This story describes the experiences and perceptions of teachers, administrators, and parents in one elementary school over the course of the school year 2001-2002. As Willow district shifts its gaze from a system with multiple approaches to a common and unified concern with New York State’s current assessment-driven system, the ripple effect is complex and multileveled. From the teachers, parents, and administrators’ perspective, however, the stakes have changed. Their performance, as measured by the state tests, has consequences both district- and state- wide. Willow Valley seems to find itself at the crossroads between teacher innovation and standardization. For example, it is not uncommon for teachers to relate that before the advent of state testing each mini-school had its own textbook or set of materials. However, the teachers also feel that under the current testing system, adopting the same textbook across all mini-schools in language arts and math is a step in the right direction. Teachers are still reluctant to give up decision-making over instruction and the unique differences inherent in their mini-school philosophies and designs, and yet the state mandated testing makes it more risky

for them to stand out alone and take a chance that their scores might not stand up to the rest of the 4th grade scores. What they seek, it seems, is support and guidance, a message from administration that makes it clear one way or the other. One that conveys an awareness and empathy for the complexity of the classroom environment, provides clear guidance on how to tackle the new state demands, as well as communicates a shared responsibility for the outcome of those decisions. Riding the pendulum alone is fine when the message supports alternative approaches to education. It results in stress and frustration when the message is one of uniqueness but the outcome needs to be the same for all.