

## **INFORMATION TO USERS**

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600

**UMI<sup>®</sup>**



**Policy in Action: The Influence of Mandated Early Reading Assessment  
on Teachers' Thinking and Practice**

**Nancy A. Place**

**A dissertation submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy**

**University of Washington**

**2000**

**College of Education**

UMI Number: 9995424

Copyright 2000 by  
Place, Nancy A.

All rights reserved.

UMI<sup>®</sup>

---

UMI Microform 9995424

Copyright 2001 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

---

Bell & Howell Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

Copyright 2000

Nancy A. Place

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Bell and Howell Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or printed copies of the manuscript made from microform."

Signature Nancy A. Pease  
Date 12/13/00

University of Washington  
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

**Nancy A. Place**

and have found that it is complete and satisfactory in all respects,  
and that any and all revisions required by the final  
examining committee have been made.

Chair of Supervisory Committee:

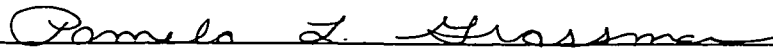


\_\_\_\_\_  
**Sheila Valencia**

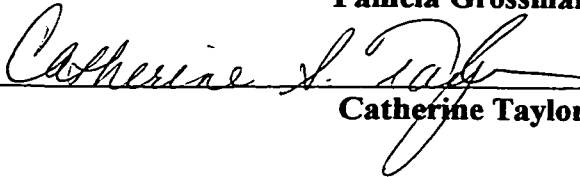
Reading Committee:



\_\_\_\_\_  
**Sheila Valencia**



\_\_\_\_\_  
**Pamela Grossman**



\_\_\_\_\_  
**Catherine Taylor**

Date: Dec 11, 2000

**University of Washington**

**Abstract**

**Policy in Action: The Influence of Mandated Early Reading Assessment on  
Teachers' Thinking and Practice**

**Nancy A. Place**

**Chairperson of the Supervisory Committee:  
Professor Sheila Valencia  
College of Education**

This study investigates the impact of a Washington State assessment policy on classroom teachers. In 1997 Washington State mandated an individually administered low-stakes second-grade reading test. The purpose of the assessment policy was to help schools direct resources to struggling students and to educate teachers about dimensions of the reading process. Building on research in policy and assessment which suggests that the implementation of state policy is influenced by district, school and individual factors, this qualitative study examines the ways in which teachers interpreted and used the results of the assessment, as well as the ways in which districts and schools mediated the impact of the state policy for teachers.

Data were collected using think-alouds, interviews and observations from two teachers and administrators in each of four schools in two districts. Analysis indicated that teachers learned from the assessment if it met an instructional need and was consistent with their theories of teaching reading. Teachers also learned from the test if they experienced dissonance between their expectations and actual student performance on the test. Teachers' beliefs and knowledge about literacy and assessment were

reflected in the ways that they interpreted and modified the test. Their interpretations of the assessment were influenced by district accountability policies as well as school contexts. Many of the teachers in this study valued the assessment because it gave them authoritative evidence of the progress of their students, whether or not they were in a high-stakes district environment.

The results of this study suggest that the Washington State second-grade reading assessment was successful in its goal of helping struggling readers obtain the assistance that they needed, while also providing useful information to teachers. Many factors contributed to the success of the policy including the fit between teachers' beliefs and the assessment, local accountability, supportive school contexts, and the nature of the assessment instruments themselves. This study suggests that it is possible to craft a state assessment policy that teachers can learn from and support.

## Table of Contents

List of Tables.....	iv
Chapter1 – Introduction .....	1
Background to the Study .....	4
Context for the Study .....	6
Implications .....	9
Research Questions .....	10
Organization of Dissertation .....	11
Chapter 2 – Review of the Literature .....	12
Impact of Performance Assessment on Teachers’ Instruction .....	12
The Mediating Role of Districts and Schools on State Policy .....	17
Teacher Change.....	21
Summary .....	26
Chapter 3 – Research Design .....	28
Study Sites.....	29
Participants .....	32
Methodology .....	34
Data Analysis .....	37
Summary .....	40
Chapter 4 – Districts and Schools .....	42
Oceanview School District.....	42
Valley View School .....	59

Oak Park School.....	70
Oceanview District, Valley View and Oak Park: Variations on a Theme .....	82
Millersville School District.....	84
Lincoln School .....	97
Webster School .....	106
Millersville District, Lincoln and Webster Schools: Small and Independent .....	116
Oceanview and Millersville Districts – A Cross-Case Comparison .....	117
<b>Chapter 5 – The Teachers</b>	
Learning From and Using the Second-Grade Test.....	123
Modifying the Test – The Influence of Teachers’ Knowledge and Beliefs	137
District Factors .....	145
School Factors .....	150
The Value of the Second-Grade Test .....	166
Summary .....	172
<b>Chapter 6 – Discussion.....</b>	<b>174</b>
Findings.....	174
Implications for Practice and Research.....	183
Limitations to the Study .....	186
Conclusion.....	189
References .....	190
Appendix A: Teacher Think Aloud .....	200

Appendix B: Semi-Structured Interview for Teachers.....	202
Appendix C: Teacher Interview Following Classroom Observation.....	205
Appendix D: Semi-Structured Interview for District Personnel .....	206
Appendix E: Coding Scheme .....	209

## List of Tables

Number	Page
1. District Comparison Data.....	29
2. Demographic Data for Schools .....	31
3. Participants .....	34

## **Acknowledgements**

Many people have helped me complete my doctoral studies and write this dissertation. In particular, I would like to thank the members of my supervisory committee, Professors Sheila Valencia, Pam Grossman, Cathy Taylor and Gerry Philipsen. They have each guided me with wisdom throughout my career as a “gradual student” and have been generous with their support. I hope that I will be able to guide my future students with as much insight and kindness as they have shown me.

I would also like to thank the teachers and administrators who took a chance and welcomed me into their classrooms and schools so that I could collect the data for this project. These wonderful people gave freely of themselves, their time and their energy, and I appreciate it.

My friends and colleagues in the Transitions into Teaching project, Susan Martin and Clarissa Thompson, have also been a source of strength. They have deepened my thinking, made me laugh when I was going to cry, and gifted me with their friendships. Marsha Riddle-Buly, another colleague and friend, was my lifeline to the university when I worked full time. My colleagues in Bellevue School District also deserve a vote of thanks for their on-going support, flexibility and the adjustments that many of them made to help me make it to class on time.

Finally, my family, especially my husband Tom, has been supportive and patient with this endeavor for over seven years. Tom has gone on more than one skiing trip alone because of the press of this work, and has not complained when I brought my laptop and books along on our vacations. He also made it financially possible for me to

“take a year off” at the end of my program to write my dissertation. Thank you, Tom, for your help and love.

## Chapter 1

### Introduction

Virginia Richardson, in her seminal article “Significant and Worthwhile Change in Teaching Practice” (Richardson, 1990) asked the question, “What is involved in bringing about significant and worthwhile change in teaching practice?” Although the data for her study were collected over 10 years ago, this question is still valid today as states and school districts struggle to help teachers improve their practices to help all children meet new and exacting academic standards.

Many states have attempted to answer Richardson’s question by mandating high-stakes/high-standards performance assessments. Washington is one of these states, mandating high-stakes/high-standards performance assessments in fourth, seventh, and tenth grades. These assessments are expected to serve as levers for grassroots educational reform by combining highly visible and rigorous standards with pressure to perform. Some researchers argue, however, that the pressure of any high-stakes environment, whether it is based in traditional norm-referenced tests or performance assessments, works against the development of significant and worthwhile change in teaching practice (Linn, 1998; Shepard, 2000b).

As a part of the Washington State assessment system, teachers were required, beginning in 1998, to give a reading test to their second-grade students. This represented a change in state policy. Prior to this time, students below the third-grade level had not been included in state testing mandates. However, drawing on a body of research that showed the importance of early intervention for those students who were struggling

readers, the state assessment was mandated in order to determine which students were below grade level. The intent of the legislation was that the test would help schools and teachers identify struggling readers and then direct resources and special instruction to these students. The purpose of the legislation was also to give teachers useful information about their students as readers. The policy aimed to encourage strong reading instruction in the early grades as an important precursor to the high-stakes, state-mandated, fourth-grade reading assessment.

The second-grade assessment represented several new directions in state policy, in addition to the already mentioned testing of second-graders. In the first place, the policy acknowledged the unreliable and invalid nature of standardized tests for young children (Kamii & Kamii, 1990) by requiring teachers to assess their students by listening to them read individually. When teachers listen to children read aloud they are better able to ascertain whether or not the students understand the directions and are on task, and whether the text the students are reading is at the appropriate level, increasing reliability. Listening to student read also increases the validity of the test by expanding the construct of reading from the silently read phonics and short passages, common to reading tests for this age group, to oral reading of longer, story-like texts. For teachers with 20-30 students in their classes, listening to students read aloud individually had a substantial impact, as it took teachers quite a lot of time. The expert panel convened by the state to craft the implementation of the legislation estimated that the assessment would take 10 minutes per child (Calfee, Duffy, Salinger, Valencia & Wixson, 1998).

Second, instead of just administering a standardized test, teachers had to use a variety of assessment tools to assess various dimensions of reading. The assessment tools were unfamiliar to many teachers and required new learning on the part of many teachers. As a reflection of the new learning required by many teachers, the state document that explained the assessment to teachers was 29 pages long (Botsford and Ryan, 1998).

Third, the focus of the policy was on classroom application and local accountability, rather than large-scale accountability. The tests were intended for use by teachers in reporting to parents, but were not intended for use in reporting to other outside audiences such as newspapers or the state. The policy was intended to be “educative” (Calfee et al., 1998) and “high stakes” were antithetical to this purpose (Shepard, 2000b). The three factors of individual administration, new learning for teachers and local accountability created a potential high-impact, low-stakes state policy – a policy that intruded into teachers’ classroom lives but was accompanied by little outside pressure.

The purpose of this qualitative study was to investigate the ways in which eight teachers understood and used this state assessment policy. Since state policies are filtered through district and school contexts, the study also examined the mediating effects of two districts and four schools on teachers’ perceptions and actions. In other words, what is the impact, “on the ground,” of this one state assessment policy?

The genesis for this study came from my experience as a district literacy curriculum specialist. As a classroom teacher I had experienced the stress, for the children and myself, of administering state and district-mandated tests. With a background in early childhood education, I especially opposed standardized testing for

young children. At the same I realized that children who learned to read fluently before third grade had a distinct advantage over children who did not. Given the personal and political realities of reading instruction, it was clear to me that struggling young readers did not always receive the instruction and resources they needed to succeed in the districts where I worked.

As a district literacy curriculum specialist, I supported the second-grade assessment as one way to help students get the instruction and the resources that they needed. I especially appreciated that the test was an individual performance assessment and low-stakes. I did not know, however, whether a mandated early childhood assessment would be welcomed and found helpful by teachers, and whether it would have any impact on their instruction or the children they served.

### Background to the Study

There are few studies of the actual impact of performance assessments on classroom teachers. Research suggests that professional development, the mediating effects of districts and schools, and teachers' own knowledge and belief systems, influence the effects of assessment policies on teachers' thinking and practice.

Spillane and Jennings (1996) posit that all policies provide "opportunities to learn." Several studies (Khattari, Kane & Reeve, 1995; Koretz, Stecher, Klein & McCaffrey, 1994; Shepard et al., 1996; Torrance, 1993) have shown, however, that professional development is an important factor if teachers are to learn more about their

students and their instructional practice from performance assessments. In a review of studies on assessment policies, Valencia and Wixson (2000) stated that

In studies which integrate professional development with assessment reform, results are most positive both in terms of teachers' learning and attitudes toward change, and in terms of useable assessment information. Reform without this support seems to produce more surface level change and questionable assessment practices" (p.37).

Because Washington State did not provide professional development to accompany the second grade test, districts and schools played an especially important role in the implementation of this state policy. Policy researchers (Hertert, 1996, Spillane, 1996, Standerford, 1997) have determined that capacity (resources and expertise) and compatibility (agreement between district and state policies) both play a role in the degree to which districts support a state initiative. Schools can also play an important role in teacher learning, especially in terms of the opportunities they provide for teacher collaboration (Hamilton & Richardson, 1995, Little, 1987, Rosenholtz, Bassler & Hoover-Dempsey, 1986, Smylie, 1988).

Finally, any teacher change is dependent on individual characteristics. How teachers perceive and act on policy is influenced by their prior beliefs, knowledge and dispositions (Spillane and Jennings, 1997). As McLaughlin (1987) noted:

Change is ultimately the problem of the smallest unit. At each point in the policy process, a policy is transformed as individuals interpret and respond to it. What

actually is delivered or provided under the aegis of a policy depends finally on the individual at the end of the line (p. 174).

### Context for the Study – The Washington State Second-Grade Reading Test

During the 1997 legislative session, Engrossed Substitute House Bill 2042 (ESHB 2042) was passed by the Washington State Legislature. This “primary reading bill” mandates that second grade teachers assess the reading accuracy and fluency of all of their students using an informal reading inventory. The purpose of the assessment is to “provide information to parents, teachers and school administrators on the level of acquisition of reading accuracy and fluency skills of each student at the beginning of second grade” (ESHB 2042, p.2). This assessment is not considered diagnostic but rather a screen which sets a high second-grade reading standard to determine which children are reading below grade level and therefore in need of extra support. In this sense, the legislature has envisioned the assessment as serving as a kind of “early warning system” for the high-stakes fourth-grade test.

The legislation required that the state select a “collection” of assessments from which districts could choose a test to use. In May 1998 a panel of experts was convened to determine this “collection” of tests. In selecting the tests, the panel was guided by a commitment to three concepts: “(1) attention to the classroom consequences of the policy (consequential validity); (2) the importance of the policy as an educative tool (educative policy): and (3) concern that the assessment is a good measure of reading (content validity)” (Calfée et al., 1998, p. 2). As an outcome of these commitments, the

panel attempted to develop a transparent scoring system because they “wanted to make visible to teachers the various components of fluency (i.e. rate, accuracy, and phrasing) so that they could be more intentional in intervening with students” (Calfée et al., 1998, p. 5). Although not required by the legislation, the panel also included an optional comprehension score because they “felt that it was important to acknowledge comprehension as a critical component of effective reading and to communicate that idea clearly to teachers and parents” (p. 4). The panel also rejected a suggested one-minute timed test of reading that was much faster to administer because they “felt that students should be given an intact passage that could be read through from beginning to end, just as they do during actual reading” (p.4). The expert panel hoped that this assessment policy would serve an educative function, helping teachers develop a deeper understanding of early reading and early reading instruction.

The Washington state second grade assessment policy grew out of strong national interest in early reading development and the belief, based on a number of national reports (see for example Snow, Burns & Griffin, 1998) that early identification and good remedial instruction for young struggling readers was the best way to insure students’ later school success. What counts as “good” instruction, however, is still a matter of controversy and debate in the public arena. The Washington second-grade assessment policy mirrored that tension when the expert committee expanded the components of the test to include comprehension as a dimension of early reading.

After the panel had chosen a “collection of tests,” districts each selected a test that they wanted to use from among these five choices. All of the tests are performance

assessments in that they require students to read aloud from a second-grade passage.

Teachers evaluate their students' oral reading on a common state rubric using the three attributes of accuracy, fluency (rate) and phrasing (smoothness and expressiveness) with a fourth attribute, comprehension, as measured through retelling, being optional and used at each district's discretion.

These reading assessments differ from others usually mandated at a state level in two ways. First, the tests require that teachers assess children individually instead of administering a standardized test to all of the class at the same time. This means that the tests take more time to administer, that teachers have to make some provision for the rest of their students while they are administering the test, and that teachers have to learn how to give the test. Teachers must also listen, judge, and score which means that they need to make decisions that reflect knowledge about reading. Secondly, results are not reported to the state or media. This eliminates public comparison between schools and districts (and actually because the districts are using different tests comparisons are difficult), thereby decreasing the level of "stakes." The legislation does require, however, that individual results be reported to all second grade parents and that a specific instructional plan be developed for those children reading "substantially below grade level." Children reading below grade level are also to be tested at the end of the year to measure their progress. Although the second grade test itself is low stakes, it is administered in a climate of high-stakes systematic reform as represented by new state learning outcomes (Essential Academic Learning Requirements or EALRs) and new state

assessments (Washington Assessment of Student Learning or WASL) at 4th, 7th, and 10th grades.

### Implications

The Washington State second-grade reading assessment policy promoted a vision of beginning reading as a complex, individual act. As a result of this vision, it placed heavy demands on teachers by requiring individual assessments of students, which were new for some teachers. Because teachers evaluated and scored each student's work, teachers' judgment was also central to the process. Although a low-stakes assessment, this policy had a high impact on teachers' time and classroom activities throughout the duration of the test, and held them accountable to parents in a new way.

Large-scale tests have become levers for reform in many states for several reasons: they are relatively inexpensive, they can be externally mandated, they can be implemented quickly, and they are visible (Linn, 1998). However, the "unintended negative effects of the consequences of the high stakes accountability often outweigh the intended positive effects" (Linn, 1998, p.33). The Washington state second grade reading test is a high-impact, low-stakes test with an educative purpose. If this test positively influences teacher thinking and practice without the negative consequences usually associated with high stakes tests, it could support the further development of a revised vision of accountability, one in which teachers and schools are mostly accountable and responsive to a local audience. In this best-case scenario, the test would help teachers

make “knowledge-based decisions” about their students (Darling-Hammond, 1989) and teachers would therefore be more instructionally responsive to the needs of their students.

Teachers are the single most important school factor in children’s academic success. All state and district instructional reforms are ultimately dependent upon the effectiveness of the teacher in implementing them. In an era of intensive school reform and high stakes assessments, research about the effects of low-stakes assessment could impact how we conceptualize “measurement-driven instruction” (Airasian, 1988) and the stakes that should be associated with test scores, with a resulting positive impact on classroom practices and student learning. The ways that teachers use and understand assessments, as well as the ways in which districts and schools support teacher learning, is greatly needed if we are to understand how policy actually plays out for teachers and students.

### Research Questions

This study examined the ways that teachers understood and used the Washington State second-grade reading test. It investigated how they felt about the assessment and how the district and school contexts within which they worked influenced their interpretations. The specific questions that guided this study were:

1. How do second-grade teachers understand and use the Washington State second-grade reading test? What do they learn from it?
2. What are teachers’ understandings of the purpose of this test? How do they react to it? How do they value it?

3. How do districts and schools mediate the state assessment policy and thus influence teachers' opportunities to learn about beginning reading acquisition, instruction and assessment?

### Organization of the Dissertation

In the chapters that follow I will present the case studies of eight teachers as viewed in the context of their schools and districts. In this chapter, I provided an introduction to the purpose and the context of the study. In chapter 2, I review the relevant literature pertaining to performance assessment, policy and teacher change. Chapter 3 is an account of the methods used in this study. In chapter 4, I describe the two districts and four schools that provided the working context for the teachers in this study. Chapter 5 consists of a cross-case analysis of the eight teachers, with a focus on what they learned from the second-grade assessment and the factors that influenced their learning. In chapter 6, I discuss the conclusions and implications of the study.

## Chapter 2

### Review of the Literature

“Policy effects are indirect and operate through and within existing settings” but in the end “change is ultimately a problem of the smallest unit” (McLaughlin, 1987, pp. 174-175). In the educational policy arena, these “existing settings” include states, districts and schools, with the “individual unit” comprised of individual teachers. In this layered study there were three strands of research that illuminated how teachers might interpret and use data from the second-grade reading test. Research relating to the impact of performance assessments on teacher understanding and instruction provide one source of information. Second are studies from the policy arena that highlight the role of schools and districts in the implementation of state mandated reform provide information about the influence of local contexts. Finally, research on teacher change provides insights on factors that might influence teacher response to this new assessment.

#### The Impact of Performance Assessments on Teachers' Instruction

The use of mandated performance tests at the state level represented a trend by states to lever as well as to measure school reform. Resnick and Resnick espoused the idea that, if “you get what you assess” and “you don’t get what you don’t assess,” then you should “build assessments towards which you want educators to teach” (Resnick & Resnick, 1991, p. 59). The assessments referred to by Resnick and Resnick were performance assessments, a type of assessment that uses direct evaluations of performances that are in themselves important learning goals, rather than indirect

indicators of competence as found on norm-referenced standardized tests. These authors argued that clear targets combined with pressure on schools in the form of test accountability could drive school reform. As a result, most state-mandated performance assessments come with high stakes.

There is little empirical evidence about the effects of high-stakes performance-based assessments on teachers' practice (Firestone, Mayrowetz & Fairman, 1998; Shepard et al., 1996; Mehrens, 1998; Valencia & Wixson, in press). However, the available research suggests that when combined with moderately high stakes, performance assessments, just like norm-referenced standardized tests, generally focus instructional activity on the test itself (Firestone, Mayrowetz & Fairman, 1998). This promotes "teaching to the test," a broad term which can apply to teaching test-taking skills, teaching content that is on the test, and teaching to the form of the test. High-stakes performance assessments are less successful in changing fundamental instructional practices, however (Firestone, Mayrowetz, Fairman, 1998). Indeed, Stecher and his colleagues (Stecher, Barron, Chun & Ross, 2000) found that the high-stakes environment in Washington State, while promoting curriculum alignment and changes in instructional strategies to match the WASL, is also encouraging teachers to "narrow their focus to the tests and to the test scoring criteria rather than the domains the tests were designed to assess" (p.63). This study also found evidence suggesting that teachers are shifting their instruction away from subjects not included on the tests, and using test preparation activities that decrease the validity of the WASL.

The evidence suggests that there are three factors that influence the effectiveness of high-stakes performance assessments in improving instruction. These three factors are (a) the tension between providing high-quality data for accountability purposes and providing data for the improvement of teaching and learning (Koretz et al, 1994; Khattri et al., 1995; Shepard et al., 1996; Taylor, 1994; Torrance, 1993), (b) professional development, and (c) local conditions.

When achievement data is used for both high-stakes accountability and instructional purposes, there is tension for classroom teachers and school administrators between using the tests to get the best student scores possible and using the tests to learn more about students (Taylor, 1994). The Washington State second grade assessment eliminated public reporting and “outside” accountability in order to focus teachers’ attention on learning from the test. It also reflected a revised definition of accountability (Darling-Hammond, 1989; Moss, 1994; Wiggins, 1993), one in which teachers and schools were accountable primarily to the local audience -- students and parents. The second-grade reading assessment exists, however, in the high-stakes climate of the 4th grade Washington Assessment of Student Learning. Whether or not teachers perceive the second grade assessment as “low stakes” is an open question and a key question, because the power of tests to affect individuals or institutions is a perceptual phenomenon (Madeus, 1988).

The administration of the test to individual students, and the use of different tests by different districts, traded off the uniformity necessary to achieve the high standards of reliability required for any large-scale high-stakes assessment with the potential for

teacher learning inherent in classroom-based assessments. However, researchers have pointed out that many teachers do not understand how to make classroom assessments a part of the on-going learning process of students, as opposed to simply serving as end-point evaluations (Shepard, 2000a).

Professional development is a factor in the effectiveness of performance assessments to change instruction because of the “ambiguous relationship between assessment, curriculum and teaching” (Torrance, 1993). In several studies (Khattri, Kane & Reeve, 1995; Koretz, Stecher, Klein & McCaffrey, 1994; Shepard et al., 1996; Torrance, 1993) a lack of clarity on the part of the teachers about performance targets was an issue particularly in terms of understanding the domains of knowledge, effective approaches for teaching to the assessments and criteria for student performance. Even though some researchers believe that an advantage of performance assessments is their transparency, or their close resemblance to real learning targets, these targets were not readily apparent to teachers in the studies cited above. Therefore, it seems that in order for performance assessments to be effective in influencing teacher practice they must be accompanied by professional development for teachers.

Three studies (Aschbacher, 1994; Falk & Ort, 1998; Sheingold, Heller & Paulukonis, 1994) looked at professional development as an outcome when teachers collaboratively developed performance assessments in K-12 language arts, social studies, science and math. All three of the studies reported significant professional growth and cited teacher ownership of the process as a crucial factor in teachers’ learning about subject matter knowledge, student performance, and assessments. Two of the studies

(Aschbacher, 1994; Sheingold, Heller & Paulukonis, 1994) also reported that teachers tended to think in terms of classroom activities and not in terms of specific learning goals, and that the articulation of criteria for judging student work for them was very difficult. All of these studies were based on teacher report and the observation of teachers in workshop settings.

However, “fundamental and conceptual change occurs slowly” as Shepard and her colleagues (Shepard et al. 1996, p. 16) found when they provided a year of staff development to teachers in their classrooms in the area of performance assessment. Similarly, in a report on the Vermont Portfolio Assessment Program, Koretz and his colleagues (1994) stated, “Despite intensive training and the provision of sample tasks, the program has not been fully able to create a consistent understanding of the desired types of practice”(p.13). Spillane and Zeuli (1999) found that instruction was a “multidimensional practice,” with “behavioral regularities of instruction” more responsive to reform than the “epistemological regularities of instruction.” These studies demonstrate that there is no quick and direct line, as some policy-makers assume, between assessment and improved student learning (Airasian, 1988) even when teachers feel ownership or when the stakes are high.

Washington State provided little professional development about the second grade test. Primarily, the state provided a workshop for district and Educational Service District leaders who then taught teachers how to administer the test. These workshops were oriented to the logistics of the administration of the test rather than conceptual understanding or implications for instruction. Ancillary staff development in the area of

early literacy was provided by the legislation in the form of competitive grants to low performing school districts. Therefore, any staff development directly related to this test had to be provided by the district or school, increasing the mediating role of these levels.

### The Mediating Role of Districts and Schools on State Policy

Spillane and Thompson (1997) argue that a district or school's capacity to carry out ambitious instructional reform depends on the ability of leaders to learn new ideas from the policy or outside sources, and then help staff inside the district learn these ideas. Borrowing from an economics model, they define the capacity of organizations to do this as "human capacity (knowledge, skills and dispositions of leaders within the district), social capital (social links within and outside of the district, together with the norms and trust to support open communication via these links) and financial resources (as allocated to staffing, time and materials)" (p.199). How do districts and schools utilize their capacity to provide opportunities for teachers to learn about and enact new reforms? Conversely, how do districts and schools utilize these resources in ways that inhibit teachers' opportunities to learn about and implement reform? The following sections highlight relevant results from past research.

#### District Effects

As a part of a large study on local perspectives on state reform (30 districts in 9 states), Hertert (1996) found that nearly twenty percent of the districts were "innovative, high-capacity districts" and generally supportive, or at least "noncritical" of state

initiatives. District policies were either legitimized by state reform or at least “not hampered.” These districts either had access to policymakers and influenced state-level policy (“assertive districts”) or relied on their own technical expertise and relationships with outside organizations (as well as their own financial stability and high student performance) so that they continued to pursue their own local agendas (“insulated districts”). Hertert found, however, that more than 80% of the 30 districts studied lacked “substantial technical expertise and capacity, access to state policy makers, and/or sufficient local fiscal resources to insulate themselves from (policy) effects” (p. 384). Hertert called these districts “reactive districts” in that state policies had a significant impact on their priorities and daily operation. These, frequently rural, districts often saw state policy as the result of political processes that favored large districts and moneyed interests and which were a poor fit with their local district needs. These districts also lacked capacity in terms of personnel, information and time to implement initiatives resulting in superficial adoptions, “picking and choosing” among state reforms and a feeling of “this too shall pass.” In addition, Hertert found that teachers’ and administrators’ ability to utilize different outside networks was a function of district capacity. Hertert concluded that

The higher the capacity and the more compatible the state reforms, the more likely it is that a district is actively implementing reforms consistent with state initiatives and the more likely it is that its personnel feel positive about the content and direction of state reform activities (p.396)

Spillane's study (1996) of two large school districts (one urban and one suburban) in Michigan elaborated on Hertert's thesis. These two "high capacity" districts increased their own reading policy initiatives as state reading policies proliferated. In one district, innovative administrators used the state reform to legitimize changes in the district's instructional policy; in the other district, administrators did not support the direction of the reform. As a result, one district's policy initiatives (and accompanying staff development and materials) supported and actually went beyond the state's reform policy while the other district's initiatives undermined state policy and simply reinforced existing practices. Both of these districts responded to the state policy by stepping up their own policy-making (so in a sense these districts were "reactive" as well). They both drew on a variety of sources, not just the state policy, to create local policy to meet local needs in the context of state reform.

Standerford's study (1997) of two small districts, one rural and one suburban, focused on the other side of the policy coin -- small districts such as those discussed by Hertert. Standerford found that these two small districts tended to respond to new state reading policies in simplified and symbolic ways that fit with what teachers were already doing so that disruptions were minimized, resources were conserved, and audiences were satisfied. Because they were smaller districts, lacking in technical and financial resources, they did not proliferate district policies, as did the larger districts studied by Spillane, but continued to follow their traditional policies and timetables in order to maintain the teaching of reading as "business as usual."

The studies cited above seem to confirm Hertert's thesis that capacity and compatibility both play a role in the degree to which a district will actively implement reforms consistent with a state initiative. Districts are composed of schools, however, and many studies point to the important role that school climate plays in the opportunities for teachers to learn from new policies and reform initiatives.

### School Climate

In studies of professional development researchers have found that a school climate that fosters teacher collaboration can play a positive role in teacher development. Smylie (1988) observed and interviewed teachers before and after an "enhancement" staff development process. (This was voluntary staff development outside of any district or school initiatives.) He suggested three antecedents to teacher change: individual psychological traits, the immediate task environment, and the interactive contexts of schools. Smylie found that, in this type of staff development, an individual sense of efficacy was significantly related to change, but that the only school context variable that was significant in producing change was interactions with colleagues. In a similar vein, Hamilton and Richardson (1995) found that the effectiveness of a staff development program may be related to school norms which encourage teachers to discuss their beliefs and practices. This finding was supported by Little (1987) who saw school norms of collegiality and experimentation as important for continued teacher learning, and by Rosenholtz, Bassler and Hoover-Dempsey (1986) who found that a "collaborative

normative climate” is connected with teacher learning while school norms which emphasize teacher individualism are not.

School norms of collaboration seem to play a role in fostering teachers’ motivation to learn from staff development opportunities. However, the picture is more complex than this. It is hard to separate school effects from district policy on the one hand and individual personality traits on the other. For example, district policy and resources can hinder or support school collaboration, while teachers who prefer collaboration could be attracted to a school that promotes this. It’s also possible that teachers may be collaborators in a school system that is either neutral or hindering. Finally, the studies cited above focused on professional development which involved voluntary participation on the part of teachers as opposed to mandated participation.

### Teacher Change

Although evidence suggests that policy initiatives around assessment will not be effective at improving teaching without professional development opportunities for teachers, there is a substantive amount of evidence that suggests that how teachers understand and implement policy or staff development is shaped by their beliefs, knowledge and disposition to learn (Spillane and Jennings, 1997).

### Beliefs

Beliefs are considered by many researchers to be an important factor in teacher change. Beliefs are distinguished from knowledge in that knowledge requires some type

of warrant (Pajares, 1992; Richardson 1994). Smylie (1988) concluded that in the context of (voluntary) “enhancement” staff development, “teachers’ perceptions and beliefs are the most significant predictors of individual change” (pg. 23) while Munby (1984) found that the beliefs of a teacher play a major role in the choices which are made about adopting research findings, implementing new curriculum or changing practice. Richardson, Anders, Tidwell and Lloyd (1991) found that there was a strong relationship between teachers’ stated beliefs about the reading process and their classroom practice, and they suggested that a lack of correlation between beliefs and practices might indicate that the teacher was going through a process of change. None of these studies, however, looked at teachers’ beliefs in the context of change initiated by institutional policies.

Just as the match between current practices and beliefs at the district level impacts the opportunities districts provide for teachers to learn, some researchers have found evidence that the implementation of new policies by individual teachers is related to the match between the assumptions which underlie the policy and the teacher’s own belief system or theoretical orientation. Cohen (1990) found that although a teacher had made changes in materials and behavior consistent with reform, her prior beliefs about mathematics prevented her from implementing the policy as it was intended. Spillane and Jennings (1997) also found that it was important to look below the surface features of reform to teachers’ discourse and the instructional tasks that they offered to students, in order to see how teachers’ beliefs impacted their implementation of the reform.

Teachers’ have deep-seated instructional patterns and beliefs that are ingrained and difficult to change (Cuban, 1993). In particular, teachers’ beliefs about testing play a

role in how they understand and use assessments. In a project that focused on classroom-based performance assessments, Shepard (2000a) found that teachers held strong beliefs about the separate nature of assessment and instruction, the importance of uniform administration of tasks, and the importance of “objectivity.” These beliefs were based in teachers’ deeply rooted positivist views of assessment and ran counter to the purposes of the classroom-based assessments, which were designed to support teachers in the ongoing instruction of their students.

### Knowledge

Another influence on teachers’ interpretation of policy is the knowledge base that they bring to the topic of the policy. This is particularly true in the case of reading policy, because teaching children to read is such a complex task. As one recent report states:

Today’s teachers must understand a great deal about how children develop and learn, what they know, and what they can do. Teachers must know and be able to apply a variety of teaching techniques to meet the individual needs of students. They must be able to identify students’ strengths and weaknesses and plan instructional programs that help students make progress. In addition to this expertise in pedagogy, teachers must master and integrate content knowledge that underlies the various subjects in the children’s curriculum. (Snow, Burns & Griffen, 1998, p. 279).

However, many teachers lack this in-depth knowledge. In addition, there is evidence that teachers’ background knowledge about content areas and how to teach in these content areas impacts their ability to understand and use the professional

development associated with specific subject areas. For example, researchers in one study (Richardson, Anders, Tidwell & Lloyd, 1991) found that teachers' understandings of the reading process played a key role in the effective implementation of a professional development initiative around activation of readers' background knowledge. Teachers who did not understand the supporting theory behind this practice were "weak and ineffectual" in their attempts to implement it. In mathematics, researchers (see for example Heaton, 1992; Fennema et al, 1996; Putnam, 1992) have found that teachers' understanding and implementation of constructivist approaches to teaching mathematics is related to their knowledge of mathematics and to their knowledge about children's mathematical understandings. And in situations where policy implementation at the classroom level is not clear, teachers must rely on their own background knowledge and belief systems (Spillane & Jennings, 1997). The result is that different teachers implement the policy differently.

The Washington second-grade reading assessment requires teachers to use individual assessment tools that are new to some teachers. Teachers have to know how to record student errors using a "running record format" or "informal reading format" (Calfee et al., 1998) in order to determine accuracy of reading. In addition, teachers have to score "fluency" on the basis of the two attributes of rate and expression. In those districts that chose to use the comprehension indicator, teachers have to score a retelling (Calfee et al., 1998). Depending on a teacher's background knowledge, all of these factors may be new information to a teacher and all of them are subject to screening through teachers' existing understandings. The extent to which the attributes of reading

exemplified by the assessment and the extent to which the assessment techniques themselves become a part of the teacher's on-going classroom assessment repertoire may be related to the teacher's existing beliefs and background knowledge.

### Disposition

In addition to beliefs or knowledge, teachers' understanding and implementation of a policy or staff development initiative is also related to their disposition or will (McLaughlin, 1989; Smylie, 1988). The construction of new knowledge on the part of learners depends to a large degree on learners' feeling that they are in charge of their own learning (Bereiter & Scardamalia, 1989). When teachers do not feel in control of their professional lives, they become compliant, act passively and do not reflect on their practice (Mathison, 1991; Noble & Smith, 1994). On the other hand, Valdez (1992) found that teachers who had been involved in a three-year staff development program developed a "change orientation" which included feelings of empowerment and autonomy. These teachers continued to reflect on their teaching and made deliberate and thoughtful changes in their classrooms. Osborn, Broadfoot, Abbot, Croll, and Pollard (1992) found a relationship between teachers' willingness to change and their beliefs about professionalism, especially as related to autonomy, the intensification of work, and changing priorities and roles. Teachers' positive or negative conceptions of changes in these areas were linked, however, to their ideological agreement with the new policies.

Teachers' feelings of autonomy have also been found to relate to their interpretation and use of assessments. Stephens and her colleagues (1995) found that

when teachers were responding to external pressure, tests did drive instruction and that the results were frequently negative. In this case teaching became a “defensive act,” superficially meeting the observable demands of curriculum and becoming a “comfortable form of pedagogy” as students became proficient at passing tests by mastering the tradition of past tests (Madeus, 1988, p.40). These researchers also found that when teachers were responding to individual learners and decisions were based on the individual or collective decisions of teachers, the test did not drive instruction in these ways.

The alignment of teachers’ beliefs with a new policy, the background knowledge that they bring to the topic area and their feelings of autonomy and control all contribute to the ways in which teachers interpret and implement a new policy.

### Summary

There are three strands of research that contribute to an understanding of how teachers might interpret and use the second-grade reading test. The first strand of research relates to the impact of performance assessment on teachers’ practice. This research indicates that high-stakes performance tests encourage changes in teachers’ instruction to match the content and form of the test, but do not encourage changes in teachers’ practice at the more fundamental level of teaching for understanding. The level of accountability and professional development accompanying a performance assessment also impacts what teachers learn from it. A second strand of research relates to policy effects, particularly with regard to the role that districts and schools play in the

implementation of state reform. The compatibility of existing district policies coupled with a district's capacity (personal, social, financial) to implement them, determines how much district leaders are able to learn from the reform, and how much they are subsequently able to support teachers in learning. School collegiality also provides an important environment for teacher learning. Finally, research on teacher change indicates that deep-seated teacher beliefs, knowledge and dispositions are important factors in the ways in which teachers respond to policy initiatives as "learning opportunities." These three dimensions of change emerged as themes in this dissertation study.

### **Chapter 3**

#### **Research Design**

In this study I examine the ways in which eight teachers made sense of the Washington State second-grade reading assessment policy. I chose a qualitative design for the study because such a design provides an opportunity for an in-depth analysis of complex issues such as teacher learning (Glesne & Peshkin, 1992). A qualitative research design also supports the development of “specific understanding through the documentation of the concrete details of practice” (Erickson, 1986, p. 121) and seems especially appropriate for a study of teaching and teacher learning, which are endeavors specific to particular individuals in particular sites (Smith & Shepard, 1998). Because state policy at the classroom level is basically what teachers perceive it to be (Jennings, 1996) and because teachers perceptions are based on a variety of factors both intrinsic and extrinsic to them, including district and school environments, I used a layered, qualitative design which would provide contextual depth to the teachers’ understandings.

In this study I investigated the impact of a state assessment policy on teachers’ understandings and classroom instruction in the area of reading. I focused on the ways in which a low-stakes reading performance assessment affected teachers and the mediating role played by schools and districts. The specific questions that guided this study were:

1. How do second-grade teachers understand and use the Washington State second-grade reading assessment? What do they learn from it? What role do teachers’ knowledge, disposition and beliefs play in the interpretation and use of the assessment?

2. What are teachers' understandings of the purpose of this test? How do they react to it? How do they value it?
3. How do districts and schools mediate the state assessment policy and thus influence teachers' opportunities to learn about beginning reading acquisition, instruction and assessment?

### Study Sites

For this study I chose two districts, two schools in each district and two teachers in each school, for a total of four schools and eight teachers. I chose two urban districts with diverse ethnic and economic populations because teachers in these districts are especially struggling to help all children meet the new standards. I selected a design of two teachers in each school, and two schools in each of two districts, to illuminate school and district influences. I purposely selected a small urban district and a large urban district to provide a useful contrast in terms of scale. (See Table 1.)

Table 1 – District Comparison

District	Millersville	Oceanview
K-12 population	2,377	47,629
Number of elementary schools	3	42
Per Pupil Expenditure	5,528	7,134
Per Cent Free/Reduced Lunch	58	47
Reading Program	Scholastic	School Choice

The small urban district, which I have called Millersville, had a population of 2,377 K-12 students, with 58% of the students receiving free or reduced price lunch and 47.24% of the students coming from minority backgrounds. There were three K-5 schools in this district. The per-pupil expenditure was \$5,528. The literacy focus in the district was on teaching reading strategies and all teachers in the district received extensive inservice training in the “Students Achieving Academic Learning” (SAIL) program (Bergman, 1992). The Scholastic series was chosen during the 1998-99 school year as a basal text adoption.

The two schools that were a part of the study from Millersville were selected on the basis of teacher volunteers. Since the research design called for two teachers at each school, the schools chosen for the study each had two teachers who volunteered for the study. Both of the schools in Millersville had diverse student populations and were about the same size. (See Table 2.)

Lincoln School had 373 students, 52% of whom qualified for free or reduced price lunch. Minority students comprised 46% of the population. Webster School was slightly smaller, with 353 students enrolled. Sixty-four per cent of these students were minority students, and 79% of them qualified for free and reduced price lunch. Both of the schools received low initial scores on the WASL, but the scores increased substantially over the years.

Table 2- Demographic Data for Schools

District	Millersville	Millersville	Oceanview	Oceanview
School	Lincoln	Webster	Valley View	Oak Park
Number of students	373	353	335	484
Per cent minority students	46	64	87	42
Per cent free/reduced lunch	52	79	67	24
Per cent meeting 1998 WASL reading standard	45.8	57.4	48.9	63.5
Per cent meeting 1999 WASL reading standard	59.6	71.2	36.2	77.4
Per cent meeting 2000 WASL reading standard	57.9	44.4	56.1	75.7

The urban district, which I called Oceanview, had a population of 47,629 K-12 students in 42 K-5 schools, which is large by Washington State standards. This district had 59.42 % minority students with 47% of the students in the district receiving free or reduced price lunch. The per-pupil expenditure was \$7,134. In the fall of 1999, the district published new grade-level standards in all subjects which teachers were expected to use to guide their instruction. During the most recent district reading textbook adoption, schools had a choice of two different basal series or a “bookroom” (trade books used for student instruction).

The two schools chosen from this district varied in terms of their test scores, the diversity of their students and their size. One school, which I have called Valley View, was one of the lower scoring schools in the district on the WASL while the second school, which I have called Oak Park, was one of the highest scoring schools in the district. The schools were both considered, by other teachers and principals in the district, to be excellent schools. The schools both had diverse populations, but Valley View, with a school population of 335 students, had 87% “minority students” with 18% of the students qualifying for English as a Second Language (ESL) services and 67% of the them receiving free or reduced price lunch. Oak Park was larger, with 489 students, 42% of whom were minority students, 12% of whom qualified for ESL services, and with 24% qualifying for free and reduced price lunch. I selected the schools in Oceanview because I knew people who worked in the schools, which gave me some access, and because I thought that the differences in diversity and test scores would provide an interesting comparison. However, as in Millersville, these schools were also selected because two teachers at each of the schools volunteered to be part of the study.

### Participants

Eight teachers volunteered to be a part of this study. In Millersville, the district curriculum specialist sent out an email asking teachers to volunteer; in Oceanview, principals at the two schools asked for volunteers. In return for teachers’ participation I offered each of them a half-day of release time to assess their students and participate in an interview, as well as a \$50 gift certificate to the bookstore of their choice. (None of the

teachers took advantage of the half-day release.) As the year progressed, I also acknowledged the teachers' help with cards and flowers. In addition, I offered each school and district free consultation or reduced-fee professional development at the end of the study. As a result, I gave a parent workshop on literacy in one school, I provided professional development materials for a teacher workshop in another, and gave a district workshop using the Qualitative Reading Inventory (QRI), an informal reading inventory.

There were a number of similarities and differences among the teachers who volunteered for the study. (See Table 3.) All of the teachers were women. They were all Caucasian, with the exception of one woman who was of Asian descent. They varied in their years of teaching experience, with three of the teachers having 30 or more years of experience, two of the teachers having between 8 and 12 years of experience, and three teachers having between 1 and 3 years of experience. Three of the experienced teachers (teachers with from 8-35 years) had masters degrees, while two of the newest teachers entered masters degree programs during the course of the study.

There were similarities and differences in the teachers' instructional situations as well. Three of the teachers taught grades 2 and 3 in multi-age classes, while the rest of the teachers taught "straight" second-grade classes. Four of the teachers kept their own students for reading, while four of the teachers organized students for reading instruction across classes.

Table 3 -Participants

Participants	Jodie	Jean	Gloria	Sarah	Sally	Louise	Dolores	Shirley
School	Lincoln	Lincoln	Webster	Webster	Valley View	Valley View	Oak Park	Oak Park
Years experience	12	28	30	3	8	1	35	3
Education	MA	MA	BA	BA	BA	BA	MA	BA
Grades	2	2	2/3	2/3	2	2/3	2	2
Core Materials	Basal	Basal	Trade Books and Accel. Reader	Trade Books and Accel. Reader	Trade Books	Trade Books	Basal	Basal
Reading organization	In-class reading groups	In-class reading groups	Cross-class reading groups	Cross-class reading groups	Individual	Whole group	Cross-class reading groups	Cross-class reading groups

The teachers also varied in their methods of teaching: four teachers used basal readers and skill sheets in reading groups as the core of their reading programs; one teacher used trade books in a whole class setting as her primary mode of instruction; one teacher used trade books in an individual readers workshop setting as the mainstay of her program; and two teachers relied heavily on trade books used in reading groups along with the Accelerated Reader program. The teachers also used a variety of assessments that had been chosen by themselves and by their districts, to determine the reading abilities of their students. The variations in teaching experience, organization for instruction, and teaching methods provided a range of interesting contrasts for the study.

### Methodology

Data for this study was collected over the course of a school year. Because teaching is a situated endeavor which is “tied to specific events and persons within the

teacher's immediate experience" (Smith & Shepard, 1988, p. 310), teachers were observed and audio-taped as they administered the second grade test to three of their students, who represented a range of reading abilities. The ways that teachers administered the test was an indication of their understandings about the test and their stances towards the test. After the test was administered to each student, teachers were asked to "think aloud" as they scored each student's assessment. This provided situated evidence of the kinds of judgments and beliefs that teachers brought to the scoring and interpretation process. (See Appendix A for the think-aloud protocol.)

Teachers were also interviewed (using a semi-structured format) about their opinions of the second grade assessment, the assessment policy and how it influenced their thinking and instruction. These interviews all took place at the beginning of the school year after the administration of the second-grade test. (See Appendix B for this interview protocol.)

Two classroom observations and follow-up interviews also provided a set of data points that helped to link teachers' stated beliefs with their classroom practices and with the ways in which they interpreted and used the second-grade assessment. Additionally, I wanted to know if the second-grade reading test had any impact on the instruction the students were receiving. The classroom observations took place from November through May. During the classroom observations I focused on the teachers' behaviors as well as the three students whom I had seen take the second-grade test. During the follow-up interviews I asked teachers about their intent behind certain lessons as well as asking the teachers about the three students I observed. (See Appendix C for this interview

protocol.) Field notes were taken during the observations and test administrations. All interviews and think-alouds were taped and transcribed.

Near the end of the school year I summarized my classroom observations, along with my description of the teachers' interpretations and uses of the test, and presented them to the participants. The participants critiqued and corrected this descriptive summary providing a kind of "member check" (Lincoln & Guba, 1985) and another window into their thinking.

In addition to interviews and observations of teachers, semi-structured interviews were conducted with relevant central office staff (e.g. the reading curriculum specialists in Oceanview and the curriculum specialist in Millersville) as well as with the principal of each school. (See Appendix D for these interview formats.) I also attended a meeting in Oceanview, called by the union representatives and attended by district officials, in which teachers voiced their concerns about state and district assessments. These interviews and observations helped me understand the messages being communicated by the districts and schools about this particular state policy and provided further information about the environments that shaped teachers' thinking.

Because teachers' discourse -- in this case, how they explain the test to others and how they interact with their students -- is important in understanding teachers' understanding (Spillane & Jennings, 1996), I planned to observe and audio-tape teachers as they explained the test during one parent conference. However, only four parents showed up for their parent conferences out of the eight that I attended. As a result, the data from parent conferences served in the study only as a supplemental data point.

### Data Analysis

This was a multiple-case study of teachers. A case-oriented approach considers the whole as an entity “looking at configurations, associations, causes, and effects within the case” (Miles & Huberman, 1986, p. 174) and then turns to comparison across cases. It is particularly useful for understanding the meanings that participants make of a particular situation and then comparing these meanings across situations. The use of cross-case analysis enhances generalizability, deepens understanding, and helps illuminate tensions between the particular and the generalizable (Miles & Huberman, 1984).

My analysis focused on the transcribed think-alouds and teacher and administrator interviews, as well as typed notes from classroom observations. Because I did not have a complete data set for the parent conferences, this data was only used in a supplemental fashion. I transcribed all of the teacher think-alouds and most of the teacher interviews myself. Another person transcribed two teacher interviews and all of the administrator interviews. In order to insure accuracy I listened to each of these tapes and corrected the transcriptions. Field notes were also typed.

Data collection and analysis was integrated and on-going (Miles & Huberman, 1984). After each teacher observation or interview I wrote a reflective memo, which helped me to think through and across the cases, and which also helped me to prepare for my next round of observations or interviews. Midway through the data collection I began to code the data for each teacher. The teacher codes at this time were relatively general and came under several main headings:

**Background**

site, years in education, level of education, previous professional development

**Beliefs**

favorite and most challenging aspects of teaching, beliefs about teaching reading

**Practice**

materials used, instructional organization, method of instruction, influences on instruction

**Assessments used in class**

district-mandated, teacher-selected

**How teachers perceived the test****How teachers used the second-grade test****Use of the test in the school context****What surprised teachers about the test****How teachers modified the test****Instruction of students below level****Teachers' concerns about the test****Professional development**

I used the codes to create a variety of matrices for the teachers. This provided a way to compare participants across similar dimensions (Miles & Huberman, 1984). I found it particularly useful to use the codes to create data displays about teachers who

had very different practices (Strauss & Corbin, 1990). I used one data matrix as the basis of the “thumbnail description” that I wrote about each teacher. Near the end of the end of the study I shared these descriptions with each of the teachers and received feedback from them. This feedback provided a check on my understandings and provided more data on their thinking for the study.

During the next stage of my analysis I coded all of the available data for each district, school and teacher (Appendix F). I focused on each case separately, examining, for example, all of the data from district administrators, school administrators and teachers that applied to Oceanview district. Because it was important to me to physically handle the data, I cut up my typed transcripts and observations by meaning units as contained in a sentence or paragraph and then taped them to filing cards. I moved the cards around on the living room floor to organize data at the different levels. General categories at the district level included district background, district policies relating to reading and assessment, literacy professional development, district policies relating to the second-grade test, district office perceptions related to state reform and to the second-grade test. General categories at the school level included school background, administrators’ backgrounds (especially in regards to reading and assessment), the organization of the school for literacy instruction, school policies regarding literacy or assessment, evidence of school collaboration, administrators’ perceptions about state reform and the second-grade test (both generally and at their school), and how administrators used the second-grade test. I used this coded data to write a detailed case for each district and school. These cases formed the basis for an analysis of the

organizational factors that influenced how teachers interpreted and used the test, which became chapter four.

I also wrote a detailed case for each teacher. To write these cases I used the previously determined codes, along with an analysis of how each teacher had administered and interpreted each student's assessment. For each case I also described how the assessment appeared to have subsequently influenced each student's literacy instruction.

During the final stage I looked across the teacher cases while asking analytic questions of the data, such as why teachers learned from the assessment, why they made modifications to the test, and what were the factors that influenced their use of it. I also focused on relating patterns in the teacher data to conditions found in the school and district data. Because of the large number of participants, data matrices played an essential role in answering these questions, illuminating confirming and disconfirming evidence, and preventing over-simplification.

### Summary

The purpose of this study was to examine the effects of a state assessment policy on the thinking and instruction of classroom teachers. The study was qualitative and layered, involving eight teachers in two urban districts and four schools. Data was collected through teacher think-alouds, interviews with teachers, principals and district administrators, and classroom observations. All interviews and think-alouds were

audiotaped and transcribed, coded, and analyzed using cross-case comparison of data. In the following chapters, the results of the study are presented.

## Chapter 4

### The Districts and Schools

This chapter is devoted to describing the two districts and four schools that comprised the teaching environments for the participants in the study. The chapter emphasizes the ways in which the districts and schools implemented the state second grade reading policy, while also focusing on other aspects of the school environment which might be expected to play a role in teachers' understandings and uses of the assessment. The juxtaposition of districts and schools highlights the contrasts between them, and illuminates the factors influencing the interpretation of state policies at these levels.

#### Oceanview School District - Tension between Diversity and Standards

Oceanview School District is a large urban district in Washington State. Nearly 60% of its 48,000 students are considered "minority" students and nearly half of the students (47%) receive free or reduced price lunches. With 92 elementary, middle schools and high schools, the district has a large central office and, at the time of this study, three central office certificated staffers who were responsible for reading programs and reading professional development in the district. (Other central office staffers supported writing as well as other subject areas and assessment.)

With a large central office staff, the district had the personnel and financial resources to respond to the new state reforms. When the school board made the improvement of academic performance the district's number one priority, district

personnel responded by creating and implementing district standards for grades K-8 based on the state Essential Academic Learning Requirements (EALRs). The standards were accompanied by a district assessment system, comprised of state-mandated tests, norm-referenced standardized tests and classroom-based performance assessments (CBA's) in reading, writing, and mathematics at each grade level. At the second grade level in reading, for example, teachers administered CBA's that were a quick phonics test, a sight word test, a non-fiction reading passage with "WASL-like" questions, as well as the state second-grade reading test which, in the case of Oceanview School District, was the Developmental Reading Assessment (DRA). The DRA was the only assessment that was commercially developed, all of the other second grade classroom-based assessments had been developed in-house by the district.

#### District Standards and Assessments: Consistency Across Curriculum Diversity

One purpose for developing district standards and assessments was to provide a measure of consistency across a large diverse district in which schools and teachers used a wide range of curriculum materials and instructional programs. As a "site-based" district, schools had their own budgets and made their own decisions about instructional materials. The district supported a diversity in reading curricula during the most recent (1994-1995) textbook adoption by selecting two basal texts and a collection of trade books from which schools could choose. In addition, several schools had adopted comprehensive, prescriptive reading programs such as DISTAR, Morningside, and Success for All which focus on direct instruction.

The diversity of reading curricula in the district was also acknowledged when the district chose an outside literacy inservice provider to lead a massive staff development initiative. (All K-12 staff in all content areas were to receive at least 70 hours of literacy inservice in 3-5 years time.) The focus of the inservice program, according to a district reading specialist, was to give every Oceanview teacher “a set of basic brain and research-based strategies so that they have a repertoire of sound reading strategies ... and have some guidance on when to use them.” One of the advantages of this particular type of inservice program, according to another reading specialist, was that the strategies could be modified to fit any type of reading program so that all teachers and schools could participate, no matter what curriculum they used.

The implementation of district standards and assessments represented a sea change for the district. Prior to this implementation, instruction had been “suggested” by a district framework, whereas now all teachers in the district were expected to teach towards the standards and administer the assessments. As one central office reading curriculum specialist noted,

We have schools and teachers that have done very much what they wanted. I have a couple of schools where literally ... every teacher has a different reading program or non-program going on and so there are some people who think that if it's mandated across the district “how dare you make me!”

With assessments in a variety of subject areas and at every K-8 grade level submitted to the central office, teachers were held accountable to the district in a way that they hadn't

been before. The district's approach to standards and curriculum was summed up by one staffer who said,

We need this (standards and assessments) very much in our district ... something that is uniform across the district so that we have a way of comparing schools to schools ... now we have two things that we have continuity with over the district. One is the standards so that everyone in the district is working towards the same goals, and at the end we have some uniform assessments to see how we're reaching those goals. And where the creativity in reading comes, is in the choice of instructional methods, schools choosing their materials, etc.

#### The Second Grade Reading Test -- Accountability and Teacher Knowledge

An area especially targeted for academic improvement was reading. In 1998 the school board passed a resolution that all students would be reading at grade level by age 8, and later that year made a district goal that by the 2003-4 school year, 90% of all fourth graders would be proficient or above on the reading section of the Washington Assessment of Student Learning (WASL). During this same time period, the state mandated the second grade reading test. As one district staffer stated, "The test was state mandated, but the Board had already given us this direction." The state second grade reading test fit into the standards and assessment system that the district was developing. As a result, there was strong district support for this test and its use was expanded beyond what was required by the state.

Oceanview School District chose the DRA for its second grade test. A committee of about 25 teachers reviewed the five possible tests. According to central office personnel, the teachers wanted a “child-appealing IRI,” (Informal Reading Inventory) and in particular liked the idea of the illustrated student books that came with both the Wright Group Test and the DRA. The DRA was selected over the Wright group because many teachers in the district were already using the Wright Group books for instruction. The committee feared that the similarity in language between the various Wright Group books and the books used in the assessment would make the test too “familiar” to the students and so “skew” the results in favor of teachers using Wright Group materials. One teacher on the committee reported that the test was also chosen because it was “kind of nicely packaged, easy to use ... it had these neat little pockets. It was an organizational thing.” The DRA was *not* selected because of a match to district curriculum or outcomes, but in fact it aligned with both the district and state standards.

#### The Developmental Reading Assessment (DRA)

The Developmental Reading Assessment (DRA), was developed by Joetta Beaver in collaboration with primary classroom teachers (Celebration Press, 1998). An individually administered test designed for students in grades K to 3, the purpose of the DRA is to give teachers a “method for assessing and documenting primary students’ development as readers over time within a literature-based instructional program,” and to “enable primary teachers to observe, record, and evaluate changes in students’

performance as readers and to plan for and teach what each student needs to know” (Beaver, 1998, pp. 6-7).

Beaver created the DRA based on her work as a Reading Recovery teacher trainer. The DRA incorporates the work of Marie Clay, who developed Reading Recovery, a program of individual tutoring for at-risk first grade readers (Clay, 1993). As a part of Reading Recovery, children are assessed for entry into the program using the Observation Survey, an evaluation instrument involving 6 subtests (Clay, 1993). Texts of graduated difficulty, used as part of the Observation Survey, are considered secure and for the exclusive use of Reading Recovery teachers. (They are not available for purchase by regular classroom teachers). In addition, the Observation Survey is designed for K-1 students. Classroom teachers wanted a way to assess the reading “levels” and abilities of their students that followed the principles of Reading Recovery and they also wanted to be able to assess their students in this manner beyond the K-1 grade levels. Beaver developed the DRA in response to these concerns.

The DRA is conducted during one-on-one reading conferences. Children read specially designed booklets which are “leveled” in terms of graduated difficulty based on text features such as match of illustration to text, amount of repetition of text, and familiarity of topic to child. (See Clay, 1991 and Peterson, 1991). The booklets are colorful and look very similar to the small trade books that many young students are accustomed to reading. Although the exact format of the DRA administration varies slightly depending on the level of the child (e.g. advanced readers silently read part of the

book, beginning readers read the whole text aloud, and the teacher reads part of the text in the earliest books) the format of the assessment for all levels is similar.

The format for the state second grade assessment requires that teachers record students' reading errors while listening to students read a second grade passage. It also requires teachers to score students on accuracy, rate, fluency, and optionally, comprehension through a retelling. The state requirements omit sections of the DRA such as student predictions about the story, comprehension questions to be used in the case of inadequate retellings, and a student interview about reading preferences. Because the state assessment was originally conceived as a "screen," state requirements do not require the teacher to go "up or down" the levels to find the instructional reading level of the child.<sup>1</sup>

The DRA was initially developed, field tested and revised by classroom teachers and Reading Recovery teachers in Arlington, OH. Although the DRA Resource Guide states, "Your school or school district will need to establish levels to identify those students who are performing below proficiency" (Beaver, 1998, p.38), the proficiency level that the Arlington, OH, teachers set for second grade readers (Level 18) is currently the proficiency level being used by Washington state to determine which students are "below grade level" in reading. This level also matches the second grade designation selected for this level text by Reading Recovery (Peterson, 1993). Because the levels are assumed to connect to certain levels in school it is somewhat "norm-referenced."

---

<sup>1</sup> The instructional book level for a student is that level at which the child achieves between 90 and 95% accuracy in decoding text (Clay, 1993). This means that the book is neither too hard nor too easy; and at the right level for instruction in strategies for decoding. The instructional book level is important for teachers to know because it can help them match students to the appropriate text in their classrooms.)

### The DRA in Oceanview District

District documents stated that the DRA in Oceanview District had two major purposes:

1. To provide better information about students' progress so that teachers can plan appropriate instruction.
2. To provide a measure of year-to-year growth towards realizing the Board resolution to have all 8-year-old students reading at standard.

The district took this assessment very seriously, for example purchasing a DRA kit, which sold for \$90, for each second grade teacher. However, some of the district actions taken to meet the needs of accountability increased some teachers' perceptions that the assessment was just something to "get through" as opposed to providing an opportunity to learn more about students. For example, the teachers reported individual student scores on bubble-in scan sheets, which gave the district an efficient way to compile data but increased the time it took teachers to record results. One teacher, Dolores, when referring to the scan sheets, reported, "It's a real pain. I don't see why we have to do this. I really don't. This is over and above what we need to do." Referring to what she did with the students' scores another teacher, Shirley, said, "So then I'm going to count up all the errors and then put it on the district's paper -- all this paperwork that we have to fill out -- so I just wanted to make sure that I tested all the kids and that I had all the kids on level."

The district asked teachers to send the raw data for accuracy and rate scores to the central office, where the computations were done, rather than expecting teachers to compute the final rubric scores themselves. The district then sent a class list of students and their standings back to the school and a letter to each second grade parent with the student's rubric scores, what the rubric elements meant, and whether or not the child met the standard. The effect of these actions, while helping teachers somewhat and insuring that parents were notified of their students' standing, also served to move ownership of the test from the teachers to the district. For example, some teachers did not know whether their students had met standard until the results were returned from the district.

The DRA was introduced to teachers through a workshop that focused on the administration of the assessment as opposed to its interpretation and use for instruction according to the teachers and administrators who attended the presentation. This contributed to the perception that the assessment mainly served district accountability purposes. In addition, because teachers were not guided in understanding this new assessment they were left to develop their understandings of the test based on previous knowledge and experience with other tests, creating an opportunity for misinterpretation in scoring. This gap was evident to district reading specialists who planned to offer inservices in the fall of 2000 on how to use assessment results to plan for instruction.

### The Expansion of the DRA

During the 1999-2000 school year, the second year of the test, the use and administration of the DRA in this district was in a state of flux as the district expanded

the use of the test. Although state legislation required second grade teachers to test all of their students in the fall, only students “Substantially Below” standard needed to be tested in the spring. However, in the spring of 2000 Oceanview teachers were notified that the district had mandated DRA testing for all second graders in the spring, as well as all first graders fall and spring. Kindergarten teachers were also notified that they should test students in the spring who had met proficiency on the district’s kindergarten reading test. The purpose of the testing was to allow teachers to “show growth over the year” and was part of the district’s goal to collect longitudinal data on students. Some principals and teachers saw the expansion of the test, which came in the middle of the school year and without district funding to support release time, as an “unfunded mandate.”

Because of the district’s desire to show student growth and not simply measure who was at standard, Oceanview also expanded the way in which the test was administered. Although the intention at the state level was to provide a screen, in mid-spring of 2000 the district stated that all students should be tested to “their highest levels of success” which meant that teachers should test to the highest or lowest levels necessary to find the student’s appropriate reading level. This expansion of the test added considerable time to the administration. For many teachers the time needed to administer the test went beyond the approximately five minutes per child, as indicated in the state inservices about the test, to approximately 45 minutes for some children, according to the teachers’ reports and my observations.

The changes in the district mandates for DRA testing were stressful for teachers, especially for those new to the profession. Teachers were sent memos from the

assessment office (“goldenrod” was the color of the paper used by this department) that notified them of changes. One first-year teacher, Louise, said of these notices,

Well -- we would get something new from the district like every week. And then I became IMMUNE to the new regulations. “Wait. Stop. Stop all testing. there’s something new. Stop all testing. There’s something new.” Literally every week we’d get a gold piece of paper saying something different. “There’s mistakes in this.” I just became immune to it and I didn’t read the rest of it...My gosh!! There was something every week!! Always, something new.

The state second-grade reading assessment also dove-tailed with the district’s focus on standards and represented a shift in philosophy. With the advent of the DRA, standards, as opposed to general developmental expectations, were established at grades 1 and 2. The district’s developmental expectations aligned with the notion of individual continuous progress in learning, whereas the new grade level standards supported the idea of uniform expectations for all children within a certain time frame. In addition, the new district standards were now based on one text, as found in the DRA, as opposed to the range of possible texts offered under the previous developmental expectations. This change offered the promise of more uniformity than the previous range of texts could provide, with the trade-off that the standard became narrower and less amenable to such individual factors as a student’s background knowledge. The standards eventually selected by the district also required students to read more complex text than were previously expected.

The state mandate for a second-grade reading test fit the needs of Oceanview district for comparable data across classrooms and grade levels. As a result, the district strongly supported the test and expanded it, surpassing the state requirements for administration. The district's orientation towards the test also signaled a shift in the district's philosophy of early childhood education. These changes, along with the focus of inservices and types of reporting procedures, increased the stress for teachers and schools while also increasing the perception of teachers that the test served a primary purpose of district accountability.

#### Communicating about the Test

For district reading staffers, one of the most difficult aspects of their jobs concerned communication. As one specialist, Abbie, explained, "The most challenging part for me is the communication issues between the everyday classroom teacher and everybody else -- getting messages down through the different avenues from the superintendent all the way down to the classroom teacher." In the case of the second grade test, communication was hampered not only by an assessment system in flux, as in the expansion of the DRA requirements, but also by the size of the district and the differing messages sent by different district departments.

Consistent administration of the second grade test across second grade classrooms was made difficult by the size of the district. Not only were there many different classrooms and teachers, but it was also hard for central office staffers to get information to them. For example, all of the teachers I observed used "cut-off" accuracy and rate

figures to automatically drop students down to the next level of the test, regardless of the total rubric score. This “cut-off” strategy could easily result in an underestimation of student abilities. The “cut-off” figures were distributed as an aide to teachers at an early DRA inservice in order to help them gauge at a glance whether or not a child might be performing at proficient level. According to central office staff the cut-off figures were never intended to be used as automatic “cut” points. Yet this understanding was consistent across the teachers that I observed, even among those who did not attend a DRA training. Although they tried diligently, central office reading staff was unable to undo this misunderstanding, at least in the schools where I observed.

Other administration inconsistencies concerned prompts and the procedures for giving the test at lower reading levels. The state administration guidelines required that teachers read to the students the title of the story and then provide a one-sentence overview of it. Some teachers at the schools where I observed did not read the title or give a brief overview of the story which likely influenced some of the students’ scores on the test. At the lowest levels, the DRA required that test administrators read part of the test to the children. In one school I watched one teacher neglect to do this, which negatively impacted and invalidated a student’s results. Inconsistencies in test administration and scoring were particularly damaging if the test was to be used for accountability. Although accountability was important to the central office assessment staff, the literacy curriculum specialists acknowledged that it would take time for teachers to learn more about the test and as a result, become more consistent in administration and scoring.

The size of the district also influenced the ways that messages from the central office were received. Because the district was so large the central office staff did not personally know all the teachers. As a result, the reading curriculum specialists said that they were sometimes perceived by teachers as “The District” or “Big Brother,” a connotation referring to the curriculum specialists as people who were really out to check up on teachers. Because the reading curriculum specialists saw their jobs as “helping bridge the fear, the anxiety that goes along with change,” building trust with teachers was an important part of their jobs. Joanne, a reading curriculum specialist continued, “And knowing... that we are out there to support them, not evaluate them, that’s a big chasm to close.”

In addition, some of the district’s actions ran counter to the message that the reading curriculum specialists were trying to promote. Such district actions as requiring teachers to report their scores on bubble sheets to the district office and extending the testing requirements promoted the perception among teachers that the DRA was a high-stakes test. Although the literacy staff thought that the scores were not currently shared with anyone outside the individual school or the central office, results were used by the central office as part of a determination of which schools were in need of academic help, and they were expected to be used as a part of a determination about which children should be retained. In addition, principals, who wanted their schools to do well, were especially concerned about school accountability since 50% of their evaluations were based on improving test scores. The reading specialists said that they expected that principals were having conversations with staff around the DRA test results, just as they

did around other tests. As a result, reading curriculum specialists acknowledged that “our teachers think this is a high-stakes test,” and that the “stress level was high” around it. All of these messages ran counter to the state intent and the message that the reading curriculum staffers. One reading curriculum specialist, Joanne, said,

We want them to understand that this is an instructional tool for them, that they can use the results to drive their instruction for the kids, not just to get a number... I see that as a major role for us to make sure that we’re making that connection to classroom instruction, not whether the kids pass or not.

In addition to the state mandate, Oceanview had two purposes for asking teachers to give the DRA; accountability and to increase teachers’ knowledge of the reading abilities of their students. These two messages as delivered by the central office were at cross purposes with each other, as the accountability message promoted the reality of “high stakes” which interfered with the message of the test as a way to help teachers better understand student progress so that they could plan appropriate instruction.

#### District-Level Perceptions of the DRA

The district reading curriculum specialists were strongly supportive of the DRA as a way to help teachers think about their instruction. For example, the reading specialists felt that the DRA illuminated the fact that some children were not being assessed by their teachers on a regular basis, and that some teachers rarely listened to their children read on an individual basis. (According to the curriculum specialists, the district provided each teacher a DRA kit to promote the on-going assessment of their

students.) The district literacy staff also felt that the DRA helped teachers to understand why it was important to group children flexibly as opposed to working with the whole class, and that this understanding had created a “real shift” for some of the teachers. In addition, they felt that the dimensions of the state second-grade reading test, especially when there were disparities between dimensions (for example between accuracy and comprehension), helped teachers view children individually. District reading specialists thought that parents benefited from this assessment as well because they could “hear the same language” across the K-2 grade levels and were empowered by an understanding of the four dimensions of reading elucidated by the test (accuracy, rate, phrasing, comprehension).

It was the opinion of the district reading specialists that despite the high-stakes environment, the feeling across the district on the part of teachers and principals was generally positive towards the DRA as well. They said that they received more positive responses than negative to the test, and talked about the number of people who had called or written to say “Thank you for your support.” They felt that the teachers valued the test as “meaningful data” that could help them see reading growth and noticed that with the advent of K-2 testing (and even third grade in some schools who had chosen to use it) there was more conversation about reading among teachers at different grade levels. As Margaret said, “I tell you, K-1-2 teachers are talking.” The district reading specialists also reported that parents who were familiar with the language and results at previous grade levels promoted the cross grade conversation. Finally, the reading specialists

observed that the DRA had promoted organization for better articulation across the grade levels in several schools.

### Summary

The thrust of many of Oceanview's efforts around the state second grade reading test grew from a need on the part of the district for comparable data between schools and across grade levels as a way to provide evidence that the district was making progress on its district reading goals. As one of the district curriculum specialists, Margaret, noted, "If indeed we are taking seriously the resolution that by age 8 every child should read, then we can't wait until they're age 8 to test them. We need to know how they're progressing along the way." This specialist had at her fingertips the fact that during the previous spring 82.9% of the second graders had met standard in the spring.

Oceanview's focus on accountability in relation to new standards represented a change for teachers in the district. The result was a high-stakes environment that was new and stressful to many teachers and principals, especially when changes were made in the middle of the year and district resources were not available to support them. This situation made Oceanview's other purpose for the test, the development of teachers knowledge about reading, difficult to communicate and achieve. Nonetheless, it appeared to district-level reading specialists that teachers in general valued the test, especially in terms of the data it could provide about children's reading progress.

### Valley View -- An Oceanview School Moving Towards Unity under Pressure

Perched atop a hill and surrounded by older homes and apartments, Valley View School was an educational community for students from many nations. With 60% of the students receiving English as a Second Language support and 67% of the students receiving free or reduced-price lunch support, Valley View's students represented a broad band of ethnic and economic diversity. The school welcomed the surrounding community into the school and I saw parents, younger siblings and community volunteers, representing a range of ages and ethnicities, every time I visited. Notices to parents were displayed in a variety of languages and instructional aides were hired in four languages -- English, Spanish, Chinese and Vietnamese. So successful had the school's community outreach program been that the principal of the school was the recipient of a prestigious award for these efforts.

The school, which was designed as an open concept building, was filled with color and activity. Children's art work lined the walls, the classrooms were richly provisioned and, over low bookshelves separating classrooms, children and adults could be seen engaged in a variety of tasks. There were many adults in each of the classrooms I walked through, partly as a result of the school's community outreach program, partly as a result of a student teaching/mentoring relationship with a local university, and partly as a function of the way that Valley View chose to spend its money -- on staffing. The average class size at the school was 22.

Jill had been the principal of Valley View School for 6 years when I met her. Prior to her appointment at Valley View, she had been a principal or head teacher in

several other schools, a curriculum consultant and a classroom teacher. All 32 years of her experience as an educator, with the exception of two years in California at the beginning of her teaching career, were in Oceanview District. She was recognized in the district as an outstanding principal, and several of the teachers moved to Valley View School because she was there. One of the teachers said, “We have a great principal, you know what I mean?” while another said that the staff “adores” her. On the occasions when I visited the school, Jill was a quiet presence in the school. Many times when I observed classes or interviewed teachers, Jill would unobtrusively come through to observe or talk with a child or teacher.

For all of the above reasons, Valley View had an excellent reputation among district principals, teachers and the staff of a local university. However, the school’s scores on the 4th grade state test (WASL) were some of the lowest in the district, and dropped from 48.9% of students meeting the WASL standard in reading in 1998, to 36.2% in 1999. In response to low state scores the school embarked on a vigorous new reading initiative.

#### Valley View’s Reading Initiative

Teachers at Valley View, like teachers in Oceanview in general, taught reading in a variety of ways. Jill commented that

Over the years, my observations have been that that there were so many approaches to teaching reading going on in this building that it was like, as many approaches as there were teachers. People were using all different kinds

of curriculum, everybody was working really hard, but the test scores weren't that great and we just, seemed like we needed more of a common focus.

During the 1998-99 school year, Jill asked an administrative intern with a background as a reading specialist, to analyze all of the school data and to meet with the teachers to try to figure out "(a) what we were doing and what their philosophy was and (b) how we could come together if that's what was needed." Using their own school study and a study funded by a state business coalition on schools that were showing test gains, it was "very clear ... that one thing Valley View needed to do was to come together on a common focus (on reading)." During the '98-'99 year the school had formed study groups to look at reading research, curriculum articulation and alignment. The staff, according to Jill, realized that they needed more help. At that point the staff decided to hire a half-time literacy specialist, a person without classroom responsibilities whose job it would be to oversee the school-wide literacy professional development process. Katy, an experienced teacher who was known by many staff members and who had also been a district mentor, was selected for the job. The job was supposed to be for one year, but Katy was so successful that the staff decided that she should continue for another year. As Jill said, "When I look at the process I think, you know, we did this right."

As a result of the staff development that Katy provided at Valley View, Jill did not encourage teachers at the school to participate in the first round of the district's literacy professional development. Jill stated, "We're probably getting more service and more knowledgeable service from her (Katy) than we would from the district." Katy, for

her part, felt that the district staff development was heavy on activities and lacked a theoretical basis, and therefore was not a very good choice for the staff.

The school put many resources behind the reading initiative. Staff agreed to devote all professional development days to literacy and Jill agreed to devote two staff meetings a month to literacy study groups. As the literacy specialist, Katy led workshops and study groups, demonstrated lessons, and helped teachers with literacy professional development in whatever ways made sense to them.

#### From Congeniality to Collegiality - "Like Pulling Teeth"

Katy saw her focus as one of developing common staff agreements about literacy. At the August retreat, although "it was like pulling teeth," the staff was finally able to agree that they would all read aloud to their students each day. As one of the second grade teachers said, "We have very different trains of thought in this school." Although the staff was very congenial ("You couldn't find a nicer staff," Katy said) they were not particularly collaborative, partly because of differing instructional philosophies. Katy described a meeting where one teacher derailed a committee's work on materials purchase because she disagreed so strongly, and several of the teachers talked about how they wanted to team-teach but it just "didn't work out," although teachers did share materials and lesson plans freely. Jill also acknowledged that the school still had a long way to go in terms of unity around reading instruction when she discussed the possibility that some of the teachers might have to move to different schools because they were

unwilling to try new things. She saw her role as one of “Give it time but be real clear about expectations.”

Based on their own in-school data and an outside report which showed that an important factor in high-achieving schools was unity of purpose in reading instruction, Valley View embarked on a literacy professional development to unite the staff around a common philosophical and instructional base —just as the school had been united around a common basal series many years earlier. By discouraging staff participation in the district literacy inservices, Jill and Katy served the purpose of working towards instructional and philosophic unity with their own staff. However, this came at the expense of becoming more unified with the district community and practices.

#### A Fortuitous Match

The guiding philosophy for the literacy professional development at Valley View might best be summed up by three phrases prominently displayed in Valley View’s staff lounge: *Know the learner. Know the strategies. Know the resources.* Jill said, “Well, it’s not a program, it’s not a program. And really it’s understanding what literacy is all about and how you can promote that throughout your teaching day.” Valley View’s professional development was also aligned with previous district statements on “Essential Elements of a Reading Program.” The focus on reading process, understanding student learning, and the view that an important aspect of instruction involved matching teaching strategies and text to learners, provided a real link for teachers to the DRA. As Katy said,

“... it (the DRA) coincides with the work we’re trying to do in the building, so it’s very fortuitous.”

Jill also saw the primary use of the test to support instruction and recognized the match between Valley View’s literacy initiative and the DRA. She said,

...they (the teachers) do use the data to determine instruction. Some better than others. Katy hopefully can help them understand how to do that and I think through the guided reading --that’s why guided reading makes so much sense to teachers because they can then use that data more effectively.

In addition, Jill thought that the DRA was an accurate test, and added that the teachers did too.

And what the teachers tell me is that it’s real accurate. If they’re (the students) getting a real low score on the DRA, they’re also not performing well at all in class, so it’s not like an ITBS where a child could get a low score but actually is performing pretty well. That doesn’t happen on the DRA. And that’s because it’s performance based.”

Because the DRA was considered to be an accurate test, it was used to help make a variety of decisions at the school level including resource allocation, retention and teacher evaluation. All of the second grade teachers met with Jill to discuss the students that they were concerned about based on the fall DRA scores. This meeting was one in which resources in terms of extra people (Title 1 or bilingual resources) or after-school tutoring were distributed. The DRA also helped in the determination of which children needed to be recommended for Special Education testing, summer school or retention.

Finally, the DRA was used at the school level as a part of teacher evaluation. Following the administration of the DRA, each of the second grade teachers made a goal about which of their students they expected to be on standard at the end of the year. In part, their evaluation was based on their ability to meet these goals.

The philosophic match between the DRA and Valley View's literacy initiative, the structure of Valley View, with its open walls and many adults and a supportive principal, made administering the DRA to second graders fairly easy. All of the second-grade teachers, with the exception of the newest teacher, met with Jill at the beginning of the year to discuss how to manage the administration of the test. Since the district required that each second-grade teacher give the test to her own students, Jill described how she then made a grid and each of the adults without a classroom assignment voluntarily agreed to take over a second-grade teacher's classroom for at least an hour so that she could give the test. During the three days that I observed second-grade teachers giving the test, a retired teacher volunteer took over one class, Katy took over another time, and a part-time teacher took a class the third time. Because there were so many adults in the building second-grade teachers did not need release time.

Although the first-grade teachers had given the test informally prior to the district-mandated expansion in order to evaluate students for placement in reading groups, the requirements of an officially-mandated K-2 DRA administration (specific time period for administering the test, reporting forms, etc.) placed stress on the school. Jill commented:

Now we'll have 10 classrooms doing it. We don't have that kind of support to provide assistance. I honestly don't know how it's going to happen. ... We have

some money and we can hire some subs perhaps to cover classrooms, which we may have to do. That's the only thing I can think of. But what concerns me about that is, it's like an unfunded mandate, you know.

Valley View had a "fortuitous match" with the DRA on three levels. At one level the match between the building-initiated professional development plan and the DRA created the context for teachers to think about the test in terms of their instruction. The philosophic accord with the DRA as a performance assessment led to its use in distributing resources and contributing to decisions about students and teachers. And the organization of the school, with its many adults, allowed for an easy administration of the test at the second grade level. The district mandate for expansion of the second-grade test and its associated stipulations, disturbed the match however.

### The Pressure to Raise Test Scores

Valley View was a low scoring school and, as such, the staff felt strong pressure to raise WASL and ITBS test scores. In the opinion of the teachers, most of the pressure came from the district. Katy said, "It comes through the district, hard and heavy, that these scores the state published will go up. And schools will pay a price. Principals will have their feet held to the fire." At the K-12 district-wide meeting called by the union in November to discuss teachers' assessment concerns, Valley View had at least five (outspoken) faculty members present in a crowd of approximately 60 people --more than any other single school and enough so that the grapevine at Valley View reported that

their principal had been called by “downtown” to question the large number of faculty in attendance.

The DRA was one of a constellation of tests that was mandated by Oceanview district or the state. Katy noted that “we have so many different testing terms floating around, people are just beside themselves. The CBA’s (new Classroom-Based Assessments) are just about what got Jill. The staff was just ready to revolt.”

Compounding the need to administer a variety of tests, Jill was also concerned that all of the tests being administered might not measure the same thing.

There’s something at every grade level in this district you know. Every grade level has a test. The difficulty is that I don’t know how much the tests correlate with each other so ... How do the scores on the DRA relate to the ITBS in third grade? How does the ITBS in third grade really relate to the WASL in fourth and then back to the ITBS in the fifth grade? Are we really measuring the same kinds of things or are we just confusing parents and confusing ourselves with all these measures?”

All of the teachers I spoke with at the school, including teachers who were not in my sample, talked about the pressure that their principal was under to improve scores. One fourth grade teacher felt that the emphasis on test scores was causing teachers to teach to the test and to ignore other things that are important to children’s learning. He felt that teachers could not build on what their students knew and needed because they were so focused on the test. This teacher called the atmosphere at the beginning of the school year “schizophrenic” because while the principal was telling people that they had

to raise scores, Katy, the literacy specialist was telling people that they should base their teaching on what the students needed. When asked about this Katy replied,

They adore Jill, almost all of them. She did succumb to some pressure to come on heavy-handed in the fall. And it was the first time those teachers ever had a negative feeling toward her, ever. This was really hard for all involved. And actually I'm really glad I was there. Because I think that those who were savvy enough to know the whole story could see that what I was saying was what Jill believes but she couldn't say it.

Katy reported that the principal subsequently "backed off" and that she, Katy, saw part of her job as trying to "de-stress teachers over testing." Nonetheless, this principal retired at the end of this school year. Although the principal assured me that she was "ready" to retire, it was hard not to wonder if the pressure over test scores played a role in her decision. (The fourth-grade teacher, who had previously been a district mentor teacher and held a Ph.D. in curriculum and instruction, also left the district at the end of the year in part due to testing issues.)

The teachers and principal at Valley View attempted to improve their reading instruction with the help and hindrance of the state and district. On the one hand, the district and state provided standards and pressure, which Jill believed was the way that schools change. In response to the question: "What goes on at the state level that you pay attention to?" Jill responded:

Standards, yeah the EALRs, and the fact that it's pretty high stakes stuff ... usually in the things I've read about change, it is an outside impetus that starts it.

Otherwise you tend to just be really happy in kind of maintaining the status quo whether it's working for people or not ... people get kind of complacent and comfortable.

On the other hand, the high pressure stressed staff and several of them were vocal that "teaching to the test" interfered with "teaching to children." The internal conflict caused by the need to raise test scores created stress and contributed to retirements and attrition. The staff's ambivalence about the tests was also likely a factor in the school's low scores.

### Summary

In many ways Valley View was microcosm of the district. In a school with a linguistically and economically diverse population, the students needed a strong instructional environment if they were to succeed academically. The staff at the school held varying, often strong, beliefs about literacy practices, and the administration was actively working to help them move past congeniality to collegiality. In the school this took the form of developing common understandings around reading practice; in the district this took the form of implementing shared standards and assessments. Pressure to improve student scores fueled this movement, while at the same time creating high stress for teachers and administrators. Expansion of the test by the district added to this stress.

The activities that the school undertook to create a common instructional purpose complimented the district-selected reading test, the DRA, and were congruent with district statements about literacy practice. However, in deciding to put their resources into their own literacy professional development instead of the district's literacy

initiative, the school struck out on its own path, choosing to develop school unity based on literacy understandings over participation in the district's professional development initiative.

The DRA presented many useful opportunities for Valley View. Because teachers and administrators thought the test was valid and reliable, it was used to distribute resources. And because of the congruity between the test and the focus of the school's professional development, the potential existed for second-grade teachers to actively learn from the test.

#### Oak Park – A High-Scoring and Conflicted Oceanview School

A modern brick building set amongst small single family dwellings, Oak Park School drew children from the surrounding neighborhood and was one of the highest scoring schools in Oceanview District. Colorful student work from the kindergarten classrooms, accompanied by captions delineating the appropriate district standards, covered one side of the large, light-filled entryway, while posters of district standards as illustrated by the work of older students graced the other side. Although all the classrooms at Oak Park were self-contained, the carpeted halls were wide and the frequent scene of individual tutoring at the low tables placed along the walls.

Oak Park School was one of the largest schools in Oceanview district, with nearly 500 students and 26 certificated teachers. Forty-two per cent of the students were classified as "minority," (meaning non-Caucasian) with 12% of those children classified as ESL and 24% of the students receiving free or reduced price lunch. The school's

scores on the '98 WASL (63.5% met the state standard in reading) increased to some of the highest in the district (77.4% of the students met the state reading standard in '99), which was especially notable because the school did not house any of the district's gifted programs.

Charlene had been the principal of Oak Park for four years when I met her. Prior to coming to the school she was the principal of an alternative school in the district, and for many years ("20 years or something") before that, she was a teacher in the district. Charlene's first comment to me when I asked her about her job encapsulated what I later saw as one of the key issues for Charlene -- tension between the forces inside and outside the school.

Charlene: Where to begin. Well, if you had to draw it, it would be an upside down pyramid. Usually a school, when you look at the list of stuff, the principal comes first and all that goofy stuff, mine's reverse and everybody else is on top of me and my job is to support all that and then sometimes it's drawn like two triangles meeting each other and the one on top sometimes, the oppressive one, sometimes it's the district.

Interviewer: You mean like this? (I drew two triangles in an hour-glass shape.)

Charlene: Yeah, yeah if it's just school that's different but the whole thing gets flip-flopped if it's the big picture. And so the district up there and the federal government and everybody else is in that, is in that big --, but me just

in that little tiny point. And school being the other big triangle. And so like trying to protect the people and support the people at the same time.

Charlene described her “biggest job” as being “somebody that not only can see the writing on the wall, but (can) translate it into a language that other people can understand...” And for Charlene, the principalship required “all your wits” -- it’s like “a whole bunch of chess games going on at the same time.”

### District Standards and Mandated Assessments

Like other schools in the state and district, Oak Park was affected by standards and assessments, although because it was a high scoring school the staff did not feel district pressure so keenly. Teachers did not set expectations for individual children based on the tests, as they did at Valley View, and the school was not in the midst of school-wide reading reform prompted by test scores. The pressure to maintain scores was high enough, however, that the fourth and fifth grade teachers did shift from multi-grade to straight-grade classrooms because they felt that “with the WASL” it was “too hard” to teach both grade levels.

Charlene and Emily, the head teacher at Oak Park who was in charge of instructional issues, both supported the district standards. Emily served on the standards development committee for the district and felt that the standards were “helpful” while Charlene commented that “I would have been a better teacher if I would have had the standards... It would have really helped to know what you were supposed to teach.”

However, although Emily supported the standards, she also had concerns. She feared that “the standards have become so important that we’re losing sight of kids, that we’re teaching standards and we’re not teaching kids.” She recognized the pressure that teachers were under in dealing with kids who came to school

...not knowing where they’re going to be spending the night, not having been fed, wondering if they’re going to be abused (when they go home) – they’ve got so much on their plates. There’s just no way in the world that they’re going to be functioning in a prime academic way. And .. it doesn’t seem like there’s any room for that.

Emily reported that the response of some teachers to this pressure was to blame the inadequacy of teachers in previous grades, and that teachers’ perceived instructional competence influenced the placement of students in the succeeding grade levels. She also noted, “And I think it’s increased the number of Special Ed. referrals. Because... the teachers are saying. “Well, if they qualify and have an IEP then we don’t have to worry about the standards. And that’s very sad.”

### The Expansion of the DRA

The district expansion of DRA testing to include kindergarten through second grade created tensions at Oak Park when the school tried to implement it -- even though the expansion was an activity that school staff initially supported. On the one hand, teachers at Oak Park voluntarily gave the DRA at the end of first grade during the 1998-'99 school year to help determine reading groups for the following fall. Oak Park also

planned to give the DRA at third grade in the fall of 2000, even though it was not required, “just to see where our kids are.” Charlene and Emily basically liked the test. As Emily said, giving the test a backhanded compliment, “I think in terms of ... all the tests that we have in the district, it’s probably the least offensive.” During the 1998-1999 school year second grade teachers learned about the DRA at such a late point in the year that they had already composed their ability-based reading groups. However, after giving the test, according to Emily, the teachers began moving their kids around a little bit saying, “Whoa, this is not what I thought it was, you know this child is doing this and we need to put him in this group.” The voluntary administration of this test at first and third grades, and its use to re-organize students in reading groups, reflected its worth to teachers and administrators.

However, Charlene and Emily had concerns about the expansion of the test by the district, which in their minds changed the nature of the assessment. With a mandate for giving the DRA K-2, accountability was attached to student results in relation to district standards. In this context, both Emily and Charlene had grave concerns about the developmental appropriateness of the test, especially for the K-1 students. Emily feared

lock-stepping kids into grade level achievement... Even though the superintendent says we should think about continuous progress... Even though the district says those are not exit criteria for each grade, we have to say to parents, “Your child’s not meeting the standards for second grade. And so what does that mean?”

Emily questioned whether it was even appropriate to expect kindergartners and first graders to read. Charlene expressed her own concerns when she noted that although she

wanted all of the second graders to be reading at level 24 at the end of second grade, she didn't want any of them "proving that they were dumb" because for developmental reasons they hadn't meet the standard earlier. In addition, like Jill at Valley View, Charlene and Emily both lamented the fact that the district required all primary teachers to administer the DRA without also supplying funding for release time.

The district's communication about time and error cut-offs also influenced Charlene and Emily's perception of the assessment. To them, standardization of scoring meant that the assessment could misrepresent a child's true ability, especially in terms of the time limitations. Charlene spoke of kids who could read but at a "molasses pace" and who would "never, never be able to pass the DRA thing ... with the time... pretty silly." Emily remarked on the tension between the district's cut-point figures, which she saw as a strategy to help teachers standardize their test administration for accountability purposes, and the diagnostic use of the assessment with individual children when she said,

We've got some kids who read just fine but they can't do it in the time frame. So they're not meeting the standards because they don't meet the time frame. But when the 1st grade teachers did their testing at the end of the year there were two kids in particular that stood out. And (a first grade teacher) wrote me a note saying, "Well this is the level that they can meet in the time frame," and it was way low, "but if you don't pay attention to the time, this is the level of their reading. And the reason that they're not meeting the time frame is they stop and ask questions and talk about the pictures." And so we don't want them to be

penalized and so that's a problem with the tests. And if the test givers, if the district trusted the teachers enough to say, "This is off the clock for this particular conversation..." (meaning that the district would say that teachers could use their own judgment about whether the amount of time a child takes to read the passage is an accurate measure of his rate) but I don't think the district would, because there wouldn't be any accurate -- it'd be too random. Different teachers would say, "Well different things were off the clock," so no standardization.

Although any single test raises issues of individual reliability, the district communication about timing cut-offs exacerbated the concerns of these two administrators. In response to a question about "Anything else you want to tell me about the DRA?" Charlene answered, "It's just the communications piece we need right now because so much in our district is mythology instead of process and procedures."

#### The DRA as a Source of School-Level Data

For Charlene, being a principal involved working at the interface of the school and the outside world. She described part of her job as being "a scout, you go out and find out what you can and bring it back to the group and give them the best picture you can about the lay of the land or where the landmines are." Among the things that Charlene brought "back to the group" were grants -- lots of them. These grants, resources from outside agencies like the federal government and the district, supported Charlene in pursuing her primary goal of meeting the needs of "every single kid." When I visited Oak Park, the school was the recipient of two federal grants, which supported 66

volunteer reading and math tutors (university and community college work-study students) and a district grant which allowed Oak Park to provide an innovative summer school for their students who were struggling in reading. This kind of money was especially important to a school like Oak Park because although it had struggling and needy students, it did not receive any Title 1 funding because it had too few students on free or reduced-price lunch to qualify.

The DRA proved to be an important tool for obtaining and distributing these resources. Charlene noted, "It (the DRA) helps me see where I need to get more support and that's one of my key jobs. It helps at an organizational level, being able to decide who gets tutors and who doesn't." Charlene also planned to use DRA data to justify and evaluate new programs, such as the up-coming summer school program. She was very excited about the possibilities of using DRA data to show gains with low-achieving children, and had already begun using DRA scores in combination with WASL scores for this purpose.

Charlene also saw that the DRA could be used in combination with other assessments to create "value-added" data for the school, by which she meant that teachers could be held accountable for the growth that individual students made during the year. Charlene described her understanding of value-added data after attending a presentation at the district office by saying,

For as long as I've ever heard about schools, always, I mean forever, Socrates, I mean there's a good teacher, there's one that's not good. And now we're going to have the data and a fair way to be able to show it. Because it's based on whoever

walked in your door compared to themselves. And there are some teachers that are going to show three year's growth for kids and other ones that are going to point six if you're lucky. Now isn't that going to be something?

Charlene's saw the DRA as useful primarily as a means to obtain resources and make programmatic decisions, as well as a possible way to hold teachers accountable for their instruction. She did not see it primarily as a way for teachers to learn more about their students. At one point in the interview, speaking of the need to give the DRA to kindergarten and first grade students, she said, "And so basically we're just going to hire a sub that's had the training and do it for us.... I want the teachers to teach, not give tests." Emily, on the other hand, noted that, because the DRA is individualized, it's "used as a diagnostic tool and ... I want to keep it that way, used as a diagnostic tool and not as an achievement level."

#### Reading and Collegiality at Oak Park

Charlene's first goal when she came to Oak Park was to develop collegiality among the staff. She described the school when she arrived as a "war zone, oh baby!"

...The staff not having a clear mission together and things like SLIG money (State funding for school-based professional development) ... Instead of saying "Here's what we really need to focus on, let's go boys and girls." It was, "Let's divide up the money and all of you do your own thing" and oh, it was awful. And so when I got here I said, "Well, you know we're not going to do it that way" and so started moving it along. There were real(ly) ...

awful problems with some of the staff, real(ly) amazing things. And so that's been a big job ... turning the staff into a group that could almost work together. I mean we're getting, we're almost to where we're "us."

One of the areas where staff conflict was most evident was in the arena of reading instruction. Charlene, although she had a vision of the kind of instruction she wanted for the school, was not really active in instruction, generally leaving this to Emily. Like Jill at Valley View, when discussing reading, Emily noted, "There's a real(ly) wide range of ways that reading is taught. From basal readers ability-grouped, to whole language, to a much more eclectic approach." Charlene and Emily both agreed that the reading instruction in the school was diverse primarily by grade level. The second-grade teachers were described as the most traditional in the school, "basal bound and ditto bound" who used ability-grouping across all three second-grade classes (called "Reading Exchange") to organize for reading instruction.

Both Emily and Charlene strongly preferred a child-centered integrative approach to reading instruction, which put them in direct conflict with the second grade reading team in terms of both instruction and organization for reading. During the 1999-2000 school year, several events occurred that began to create some change in the second-grade team. At the beginning of the year, the first-grade teachers opted out of the Reading Exchange (with the "total support" of Emily) and the reading resource teacher (a part of the Reading Exchange) was given a reading group based on DRA scores as opposed to having a reading group equal in number to the reading groups of the regular second-grade classrooms. Later in the year, two of the three second-grade teachers were a part of the

school team participating in the district sponsored literacy professional development.

Both of these teachers found the inservices “wonderful” and Emily commented “...that’s why I was so glad that two of the second-grade teachers went to this literacy training.

That’s going to maybe shift things. Maybe we might even get out of Reading Exchange.”

Emily and Charlene hoped that two more activities might also change the way that the school and particularly the second-grade team, organized and taught reading. One was related to the summer school literacy program, which used a high-interest integrative approach to teach literacy as opposed to, in Emily’s words, “the same drill and kill and read out of the basal.” Emily was hopeful that, if it were successful, the summer school literacy program might serve as a “model to infiltrate the school all the way through.” Another change related to the regularly scheduled afternoon school releases, initiated during the 1999-2000 school year, initially designated for grade-level team meetings. Emily planned to “reconfigure” the groups for the 2000-2001 school year to cross-grade groupings in order to develop more coordinated instruction across grade levels, especially in the area of reading. She noted “...probably the most tense and the most important (are) the conversations that need to happen between second and third grade teachers...” Emily felt that it was a big transition for second-grade students to move from classrooms that were “very structured and very traditionally oriented and very work-sheet oriented to take the leap into a totally, totally reading-for-meaning, reading-for-purpose literature-based program.”

Although Charlene and Emily wanted to change the content and structure of the second-grade program they were at a disadvantage because their school reading scores

were so high. As Dolores, one of the second-grade teachers remarked when telling me about their school WASL scores, “We must be doing something right!” Shirley, another second-grade teacher noted that “Reading scores are good... because, if the scores weren’t up there, I think maybe we would change our program.”

### Summary

Both of the administrators at Oak Park valued the DRA -- but for different reasons. Charlene, who saw her job as one of dealing with the outside world and providing the staff with the resources they needed to “help every single kid,” valued the DRA for the data it provided. These data helped her to obtain new resources (in the form of grants) for the school, provided a way to evaluate the success of new programs, served as a mechanism for the allocation and control of resources within the school, and had the potential to measure the success of teachers’ instruction. The fact that teachers voluntarily administered the assessment beyond the mandated grade levels showed that administrators and teachers found the test helpful and encouraged its many organizational uses.

At the same time Emily, in particular, recognized the diagnostic potential of the test and how it could be hampered by the standardization called for in a district mandate. When the district mandated the DRA for K-1, after knowledgeable first grade teachers had used it voluntarily, it became obvious that teachers’ flexibility to use their own expertise in scoring and interpreting the test was compromised. The potential conflict between teachers’ knowledge about students’ reading progress and the requirements of

standardization was highlighted by the misunderstanding about cut-points for time and accuracy initially distributed by the district. In addition, the expansion of the test mid-year created stress on the schools. Charlene's plans to use the test data for direct teacher accountability, as in the value-added plan, also ran the risk of discouraging teachers from using the test to think about their students or their instruction.

Like Valley View, there was great diversity in reading instruction at Oak Park. The second grade team at Oak Park, in particular, was in conflict with the administration and some of the other teachers at the school over reading instruction and organization. The DRA, reaching across three levels and used for organizational purposes, played a role in this conflict.

#### Oceanview District, Valley View and Oak Park -- Variations on a Theme

Oceanview District had a variety of reading curricula in place across the district. The two schools in this study mirrored this variety with diversity in reading instruction among the teachers at each school. The district and both schools were searching for a way to bridge the diversity: the district by having common standards and assessments, Valley View through its literacy initiative, and Oak Park through the actions of its administrators. However, the reasons each of these institutions wanted to bridge the diversity in instruction varied. For Oceanview district, common standards and assessments were a way to be able to compare schools as part of accountability in large system. At Valley View, bridging diversity meant intensive staff development in reading to help teachers develop a common purpose and collegiality. At Oak Park, bridging

diversity meant aligning second grade teachers with the administrators' understandings of good reading instruction. In each of these situations, the state second grade reading assessment played a role. At the district level it was strongly supported as a vehicle of accountability through inservice and reporting activities. At Valley View the district-selected test matched the school's professional development orientation. And at Oak Park the test was used by both the administrators and second grade teachers to support their own views on instructional practice.

The district had two purposes for using the DRA -- accountability and the development of teacher knowledge. Because of the way that accountability strategies were implemented and because the initial focus of professional development was on test administration as opposed to interpretation and use, these two purposes were often in conflict with each other, creating some confusion and stress for teachers and schools. On the one hand, district actions increased the perception of teachers that the test was high stakes and mainly served the purpose of accountability; on the other hand, district reading specialists were trying, through inservices and personal contact, to help teachers use the test results to inform their instruction. The efforts of the reading specialists, hampered by the sheer size of the district, limited resources, and lack of personal relationships, was no match for the capacity of the district to manage and report large amounts of data. The second purpose of the test, the development of teacher knowledge, was overwhelmed by the district's need for accountability data.

It was not just the district that used the test for accountability purposes. Both of the schools used DRA data as an important measure in the distribution of resources and

the evaluation of students for such programs as Title 1, summer school and retention. Both schools also considered the DRA in terms of teacher accountability. At Valley View this accountability was in place; at Oak Park the principal was intrigued with holding teachers accountable for students' growth over the year. The principal at Oak Park also found the DRA scores useful in obtaining grants and evaluating their effects. Both of the schools had also found the DRA scores useful for grouping students for reading, and had voluntarily expanded the test to first grade before the district mandated it. The DRA filled a need at both the district and school levels for an assessment of young children which teachers and administrators considered valid.

#### Millersville School District --Small, Urban and School-Centered

With 2,377 K-12 students, 5 schools and 160 teachers, Millersville was smaller than some high schools found in urban districts. It included an urban population that was ethnically and economically diverse, with 58% of the students receiving free or reduced price lunches and 47% of the students considered to be from "minority backgrounds."

The central office staff of the district was quite small as well. A superintendent and an "assistant superintendent without the title" in charge of all categorical programs (special education, nurses, occupational therapists, etc.) made up the central office administration while Liz, a certificated teacher, served as the coordinator in charge of all K-12 curriculum programs. Liz stated that her role was "to oversee any of the curriculum and staff development work that's necessary for all subject areas, all grades. And I feel just a little overwhelmed!" She continued:

I came into this role in ... school year '94-'95 so the Commission (Commission on Student Learning) and all that work was just picking up, so it's been six years of "Can't we just -- I just want it the same. I don't want anymore cognitive dissonance -- it's totally hard!"

As Liz's comments indicate, with a small administrative staff, the central office lacked the capacity to support teachers in meeting all the new state standards and accompanying assessments. Liz felt this lack of capacity keenly, especially with regard to her most "pressing piece," the creation of district curriculum guides to match the state EALRs. She bemoaned the fact that although she had the money to support teacher development of the curriculum guides, there were only 160 teachers in the whole district to do the work, and out of that, only 20-30 teachers who would actually do it. She did not know how this task was going to be accomplished K-12 in all curriculum areas.

The small size of the district also meant that Liz knew everyone in the district personally. This was both an advantage, in terms of personal relationships, and a disadvantage, because according to one teacher, everyone expected her to do what they wanted. It meant that people thought of the central office in a personal fashion, as revealed by teacher comments such as "Oh, we just send it in to Liz" or "Don't tell Liz, but I do it differently." Although Liz was very well liked in the district one teacher noted, "I felt really badly for Liz sometimes. People would take things personally, get mad at her because she didn't do it their way." In some ways, because Liz was known personally by all of the staff in the district, and also perhaps because she was categorized

as a teacher and not administrative staff, she had less “clout” when it came to enforcing district curriculum or assessment guidelines.

### School-based Decision-Making

The Millersville School District also operated with a building-centered governance model. This orientation, coupled with a lack of capacity at the district level, created the space for schools to develop their own curriculum. Liz commented that one of her on-going questions was, “How does the central office fit in a governance structure that’s really a site-based endeavor?” She said,

We’ve been in sort of, kind of, this loosey goosey environment of giving people a lot of freedom, not having a real good sense of boundaries. We’ve tried to also, in the last several years, to say , “What is our implementation process for curriculum? What does it look like?” I mean, how do you go about doing that? We certainly have Board policy that clearly outlines it, but ...

Liz said that teachers or department chairs would buy textbooks without going through district processes. The district also gave each teacher \$750 annually for supplies, books or professional development, which many teachers used to buy books or training they felt they needed and which also supported teachers in working outside of a district curriculum. At the same time, without clear curriculum guidelines (aligned to EALRs) in many subject areas, the schools and teachers were “working in a vacuum” as they made their choices. Even when there were clear guidelines, as when the district adopted math

and reading textbooks, the schools tended to purchase their own curricula to supplement or replace the district adoptions.

Rather than seeing the schools' independence as a liability, however, the district actively cultivated it. In 1997, according to the principals, the superintendent informed them that they would each have to select a "comprehensive school reform model." Liz recalled him saying that schools needed to decide on a reform model because "we have a district where we have to ... be really smart about what we're doing. Our kids aren't going to come and make it if we don't." After teams of staff members visited a state-sponsored conference about various school reform models, and with the encouragement of the superintendent who supported it because he thought the program especially appropriate for schools with a high population of poor students, the Accelerated Schools model (Levin & Hopfenberg, 1991) was selected by all three elementary schools and the middle school. Liz said that the district didn't think it would make sense to do a curriculum focus, like Success for All (Madden, Slavik, Wasik & Dolon, 1997) or Chicago Math (<http://www.social-sciences.uchicago.edu/ucsmp>) because "...you haven't even asked the question: What are the needs? What do we need to do?" The district supported the Accelerated Schools process not only philosophically, but also by making a 3-year commitment to a local university for training and by giving financial support to schools for release time and professional development.

In many ways the Accelerated Schools model made Liz's district curriculum work more difficult. As part of the Accelerated Schools process, schools developed literacy and math cadres to improve student learning. One of the messages that Liz got very

clearly from staff was, “Don’t have curriculum committees for literacy. Every cadre’s already researching the question about literacy, why would you have a duplication? We don’t have enough people and you’re going to burn folks out.” Liz recognized her dilemma when she said, “I have so much -- I have pieces I need to get done, and I’m not in charge of this site-based stuff!”

The adoption of the Accelerated Schools model as the district’s response to state reform initiatives placed demands on the teaching force while also strengthening and legitimizing the decision-making power of the schools. As a result, it was more difficult to implement district-wide curriculum. Although principals and teachers reported that, in such a small district, individual teachers could have a lot of influence on district policy, and there was evidence to support this, adoption of the Accelerated Schools model reinforced the strength of school-based as opposed to district-based policy.

#### Assessment in Millersville

Student assessment was an increasing concern for the administrators in Millersville. One of the principals reported that there was an on-going discussion about ...what we’re going to do at each grade level because OK, we have the test at second grade, how do we know that third grade are still on line so that when they get to the fourth grade test ... we need some measures along the way so that the ones that begin to drop out, we can catch them a whole lot earlier than what we’re catching now. So yeah that’s been part of the discussion in our district you know,

does there need to be a kindergarten test of some sort so that we know that they're on line to pass the second grade one.

Until the 1997-98 school year, the CTBS (CTB/McGraw Hill, 1990) was administered from third grade through high school (according to Liz, "because we didn't know what else to do,") with no district testing mandated for earlier grades. However, in 1997 the district began using the Gates McGinitie test (a standardized, norm-referenced group administered test, MacGinitie & MacGinitie, 1989) once a year and the STAR test (a computer-based reading comprehension test designed to determine students' reading levels, [http://www.epicent.com/software/products/pages/s/star\\_reading.html](http://www.epicent.com/software/products/pages/s/star_reading.html) ) twice a year at second grade. These tests were used as "markers" of student progress and student results were reported to the district. Liz considered these "big markers" however, and felt that there was a need for on-going classroom-based assessments. She asked, "How do you know when you need to move kids or change things or what needs to happen next?"

Because of her concern about classroom-based assessment, Liz was very supportive of the state second grade reading assessment mandate. She noted that prior to this mandate, district teachers had not had any formal training in running records (a method for assessing student reading errors, Clay, 1991) or Informal Reading Inventories (IRI's) and that even for those teachers who had learned how to do IRI's in some other way, it hadn't "become a way of doing business." She liked the idea of the second grade reading test she said "... because I believed it was institutionalizing a really good classroom-based assessment tool." The second grade test filled a need for Liz in that it gave her a way to approach classroom-based assessment with the teachers.

Liz's approach to the state second grade assessment mandate was low-key and reflected the limited capacity of the district. She made the decision to use the Qualitative Reading Inventory (QRI, Leslie & Caldwell, 1995) herself, because the teaching staff was on vacation when a decision on the state legislation was reached and the timeline for implementation was very short. She based her selection of the test on pragmatic factors: a recommendation from a neighboring district that the test was "user friendly;" her disinclination to purchase any more little books (referring to the DRA); and because of the guidance she thought that the QRI provided in relationship to the retelling assessment. Liz hired a staff developer from a neighboring district to provide an after-school workshop on administering the test. She also sent teachers a packet of materials that she had collected from another district relating to the plans that they might write for their lowest achieving children. Her approach, she said, was, "Here's some ways to approach the plan. I didn't say do this or this, but here's some ways to do it."

Despite her low-key approach, Liz did have some definite ideas about test implementation. She was particularly concerned that teachers wanted substitutes to cover their classes. Although she paid for subs the first year, she did not want to send the message that assessment should happen just once a year, and she envisioned the kinds of assessments that were a part of the second grade test as being "on-going." Although she thought that most of the teachers were not doing on-going assessments modeled after the parts of the QRI, she was "really glad to hear" that one teacher at Webster was using Individual Reading Inventories to move kids between groups. Liz said,

That's what we want to see happening, ... to have the assessment being a matter of course, not some separate thing that we're required to do once, twice a year. In other words, it's informing their practice, which was the whole point.

Liz strongly supported the state-mandated second grade test but she was hampered in this area by lack of district capacity. Because of the broad scope of her job, the state second-grade test was only one area of responsibility among many. Although she recognized the need for professional development about classroom-based assessment, little professional development in this area had been done in the district. She noted that the district had offered one 10-hour reading assessment class ("sort of a study group") and commented, "We really haven't done enough in helping our folks have good classroom-based assessment tools." Although she had fully intended to talk with second grade teachers about the test during the first year, "the year got away and we didn't get a chance to do it." She also couldn't remember whether or not the district required teachers to use a retelling to test for comprehension. (The district form has a space for retelling scores.) Although teachers were required to turn in their fall and spring results to her, Liz had not looked at them and did not know, in fact, if all of the teachers had turned them in. Ideally, she said, the principals should talk to their second grade teachers about the students' results and the plans that they wrote up for the parents. "Realistically," she said, after having taken over a school for several days as principal, "You got no time to do anything!"

In Millersville, second grade teachers were required to administer the QRI and report scores to the district office along with two other tests, the Gates McGinitie and the

STAR test. Although the QRI was supported by the district curriculum specialist as a way to encourage the development of classroom-based assessment strategies, the limited ability of the district to encourage this development, coupled with the use of two other commercial tests at the second grade level, made it difficult for the curriculum specialist to communicate her vision of the QRI as a useful classroom-based assessment.

### The Qualitative Reading Inventory (QRI)

The QRI, developed by Leslie and Caldwell (1995), is an individually administered informal reading inventory (IRI) designed to “provide diagnostic information about (1) conditions under which students can identify words and comprehend text successfully, and (2) conditions that appear to result in unsuccessful word identification, decoding and/or comprehension” (Leslie & Caldwell, 1995, p.1). Although the QRI has many sections (word lists, questions for eliciting background knowledge, reading passages, retelling format and comprehension questions), teachers were only required to listen to children read one passage, and score children on their accuracy, rate and phrasing to meet the state guidelines. They did not have to administer the comprehension sections nor did they administer multiple selections as required for the QRI.

The QRI passages were written based on passages in basal readers and in content area (science and social studies) textbooks. Content and the length of the passages were based on these texts. The difficulty of the passages was estimated by average sentence length as well as vocabulary difficulty (based on word frequency and number of

syllables). Readability formulas were also used to measure the grade levels of all passages. The difficulty of all of the passages was also assessed by asking a fairly small number of children (for example, 19 students read second grade passages) to read passages at adjacent readability levels.

The QRI is not a not a norm-referenced or standardized instrument. The manual refers to the examiner as a “decision-maker” who “must decide what information s/he wants to obtain about the student and how to use the QRI-II to obtain it.” Although the state format for delivering the second grade test as a reading ability screen eliminated the need for much teacher decision-making, when the test was used to find reading levels or to obtain diagnostic information on children reading below standard, teacher knowledge became an important factor.

#### The QRI and Literacy Instruction in Millersville

As noted previously, the QRI was chosen by Millersville district for pragmatic reasons and not because of a strong match with its curriculum or instruction. However, the district’s long running staff development program, as well as one of the literacy curriculums in use district-wide, did have the potential to impact teacher’s understandings and use of the test.

The district’s longest running instructional effort was in the area of reading. For six years, the district had provided extensive long-term staff development in SAIL (Students Achieving Independent Learning, Bergman, 1992). This was a program that helped teachers learn to instruct students in 6 reading comprehension strategies -- getting

the gist, predicting, thinking aloud, visualizing, summarizing, and problem solving.

Using a rationale reminiscent of the one used by Oceanview to justify its major literacy professional development initiative, SAIL was initially selected because its instructional strategies applied across all K-12 levels and could be used with a variety of materials.

Furthermore, the program had been developed with struggling readers in a large, poor district. As Liz said, "So it fit ... the vision. It would work." All of the teachers in the district were expected to take a five-day training in SAIL in their first year, and then take additional training over the next two years. Liz explained that one of the reasons the district was able to provide this professional development "faithfully" was that she "sat through" a lot of the training and was able to give it herself. This demonstrated how the resources in this small district impacted decision-making in the area of professional development.

Liz was not sure how extensively SAIL was being implemented. The expectation was that teachers would explicitly teach strategies from SAIL at least three times a week and that teachers would understand it well enough to teach it in a way that developed "on-line processing" in students. As Liz noted, the program was not "prescriptive," there was no "set way , ... no recipe, ...no answer. It's your decision as a teacher."

Another official district curriculum was the recently adopted literacy series from Scholastic Publishers (1995). Selected by a group of district teachers and administrators, this literature-based series was chosen primarily because it provided good materials for teaching SAIL comprehension strategies. However, teachers at Lincoln, one of the schools where I observed, used the previous Harcourt Brace Jovanovich series (1984) as

their primary reading text, while teachers at Webster primarily used sets of tradebooks and Accelerated Reader as the core of their program. As supplemental materials, the Scholastic materials did not receive much use or attention from the teachers and were, therefore, a minimal element in the instructional environment surrounding the QRI.

Unlike the Scholastic series or SAIL staff development, Accelerated Reader (<http://www.advlearn.com>, commonly called AR in the district) developed as a grass roots movement in the schools. Liz described AR by saying, “That was one of those curriculum endeavors that started one place and now it’s a district program.” Lincoln School was the first to use AR, supporting its use from school funds (which it continued to do). The other schools followed, and were supported through district funds.

AR is a computerized program that, in conjunction with the computerized assessment STAR, matches children to instructional level trade books (library books). After reading a book a child takes a multiple-choice computer test that measures his or her literal comprehension of the book. The program keeps track of children’s reading using a point system based on the difficulty level of the book which they read (more difficult books were worth more points) and provides an official-looking printout each time a student takes a test so that teachers and parents have an up-to-date record of student reading and comprehension. Although the schools varied in terms of how much they used prizes to reward students for achieving points in AR (a recommended aspect of AR programs), the fifth graders across the district competed with each other for pizza parties. (The losing schools paid for the winning school’s party.) Although Liz had concerns about the “unintended consequences” of the program (the reliance on extrinsic

motivation, inflexibility in terms of the books that students could read, and the literal nature of the comprehension questions), she said, “But our teachers have said (that) the kids are really motivated you know, so it’s really a sacred cow.” She added, “And nobody is asking about the questions. Nobody has said, “What level of Bloom’s taxonomy are those questions? Is this really preparing our kids for the WASL?” Although limited to an assessment of children’s literal comprehension of the books they had read, the program’s frequent printouts of each student’s reading scores on the AR tests served in both schools as an important source of on-going classroom assessment.

Millersville’s district-wide literacy staff development and curricula created a variety of contexts within which teachers’ understandings and uses of the QRI were situated. A long-running staff development program created a set of expectations about reading instruction and the strategies that students should use when reading, while a grass-roots initiated curriculum created a context of on-going assessment and had the potential to influence teachers’ understandings of comprehension. Meanwhile, an officially adopted reading curriculum was not extensively used and therefore did not much influence the context of teacher’s instruction.

### Summary

As a small urban district, Millersville faced the challenge of meeting difficult new state standards with an economically and ethnically diverse population and limited district capacity. The district chose to meet the state standards by enhancing each school’s ability to make instructional decisions. Using the Accelerated Schools model,

the district devoted resources to strengthening school-based decision making which made it difficult to develop or support district-wide curriculum or assessment efforts.

Previously mandated tests at the second grade level, as well as the multiple tests that accompanied the use of Accelerated Reader, created a context for the QRI as one assessment among many. Although the district curriculum coordinator was highly supportive of the state second grade reading assessment mandate, and valued it particularly as a way to develop classroom-based assessment, she was limited in the attention she could pay to it. As a result of the above factors, the QRI in Millersville was a low-stakes test. There was little pressure from the district with regard to results, and there was little instruction or guidance from the district office for understanding or using it.

#### Lincoln – A School in Millersville Where Teachers Were the Experts

Lincoln School was obviously a school in the midst of change. Surrounded by the constant noise and dust of a completely new school under construction (on the playground), the school staff nonetheless provided a welcoming environment for students, parents and other visitors and seemed to accept the vagaries of construction with grace and resourcefulness. For example, I heard over the intercom after a rainy morning: “The Rooter Man is fixing the drain on the playground again and so we have to have indoor recess. I’m really sorry, but there’s nothing we can do about it.” The teacher replied, “That’s OK we’re ..” and the class answered in unison, “Flexible!” Lincoln was a diverse school with students speaking 21 different languages, 46% of its students

representing “minority” populations, and 52% of the students receiving free or reduced price lunch. Out of 373 students, only 57 had been at the school for four years, illustrating the major challenge that student mobility posed for Lincoln’s 29 full and part-time teachers. Despite all of its challenges, Lincoln’s scores on the WASL steadily climbed from 33% of students meeting standard in reading in 1997, to 46% in 1998, to 60% in 1999.

The staff I spoke with at Lincoln Elementary was unanimous in their praise of Mike, the principal of the school, and he reciprocated with strong admiration for them. Mike had been principal of Lincoln Elementary for three years, coming to the school after a number of years as a special education teacher and principal in other districts. He said that, before he came to Lincoln, he was considering a career shift out of education but that the staff at this school had “put the energy back in his step.” He described his tenure at Lincoln as being “heaven on earth, just the staff is so good.” Mike noted that he had been the third principal in three years for the school when he came, so that the staff “had had to rely on themselves” and for that reason they were a “tight staff -- in a very positive way.” Mike had come into a situation where the staff had strong ideas and were used to taking initiative.

### Support for the Staff

Mike’s view of the principalship meshed with the type of faculty he had. He saw his job to “keep them (staff) motivated, keep them pumped up, make sure that they’ve got what they need to teach the kids whether it be a curriculum item or just the energy level.”

As a result, he supported his staff in attending professional development activities of their choice, initiating proposals and developing curriculum. When the new school construction required that the school lose a second grade classroom, the second grade teachers created a proposal for a Reading Immersion Class, a small class (maximum size 15 -- meeting in a very small space) that would focus on literacy skills all day with the second grade's most struggling readers. At the same time, the other second grade classes, having to pick up larger class loads, would each have a full-time instructional aide. The second grade teachers initiated this proposal, Mike supported it, and the school board approved it for the 1999-2000 school year. The teachers called this program a "pilot program" and decided that they wanted to present data about increased student learning as a result of the program in order to persuade the board to continue it. In initiating this proposal and the subsequent data collection, the second grade teachers turned a construction mandate into a winning situation for them and for students. They also took ownership for data collection about student learning which included the QRI.

Mike supported his teachers' ideas, even when they ran counter to district ideas. For example, as a part of the Accelerated Schools process, Lincoln School had a reading cadre designed to study reading in the school. According to the teachers, the Lincoln cadre focused primarily on how the school could improve reading instruction by developing or obtaining materials with more controlled vocabulary and phonics than were a part of the Scholastic Series adopted by the district. Mike agreed with his teachers that Scholastic was problematic, especially after one of the second grade teachers made a presentation which demonstrated the variable reading levels of Scholastic materials

compared to the listed reading levels for the same materials on the Accelerated Reader. Mike's support for teachers created an environment in which the second grade teachers defined their instruction and assessment in terms of what they thought was needed, even if it went against the announced district practice.

Mike saw his staff as "tight;" even though they represented a variety of teaching styles. He gave a lot of "leeway" to the staff and said

... I tell them over and over, "We've got to get from A to B and how you get to B, I don't really care, just get them to B." Of course, I have some care along the way, but my point is, get them to B and that's great because I trust you. You're here, I've watched you long enough, I know you're good, so just get the kids there.... And I know they use that openness and that leeway because we've got a whole lot of different teaching styles and we all know that there's a whole lot of different learning styles.

So strong was Mike's belief in the importance of teachers teaching in a way that was true for them that there was some tension between teachers in the school around the variety of instructional styles. One example of tension was caused by the third-grade team who used an organizational model of ability grouping across three classes for reading, math and language arts. According to one teacher, the second and fourth-grade team did not think this type of instructional organization was good for students. They believed that this model was too rigid to meet the needs of students, and that student motivation would be diminished for those students in the lowest group. This disagreement was discussed extensively in the math cadre's meetings, because as one

participant said, “There are no third-grade teachers in there!” Although, according to this teacher, the second and fourth-grade teams asked for support in changing the third grade teachers’ instructional organization, Mike said to them, “It’s working for those guys so just leave it alone.”

### The Importance of the Individual

Mike saw personalized interactions with teachers as a key ingredient in his work with them. Mike met with each teacher once a month. The teacher brought her class roster and recent relevant assessments (such as the QRI) to each of these meetings, and talked to Mike about each student. Mike said that this process was beneficial in two ways. One, he got to know the students better and could acknowledge them or remind them about their work when he saw them around school. Second, it kept each student fresh in each teacher’s mind. Mike said he’d ask questions about students who were not successful such as, “So what are you doing differently? It doesn’t sound like that’s working.” This individualized approach carried over to the Title 1/Special Education teachers in the school who mentored the Instructional Assistants on a weekly basis in the same way.

Mike’s emphasis on the importance of each individual was a good match for the state second grade reading assessment with its requirement for one-on-one interaction between teacher and student. However, in the same way that Mike valued each teacher’s own instructional stance, Mike valued the teachers’ daily on-going assessment in relation to their own instruction more than the QRI. He said,

Well, like any test... OK that's their score for that day, that hour and that minute, ...but I always ask my staff, after they give me those scores, what's Johnny doing daily? What's his daily progress? What's his daily scores, if I can use that term? Where is he on every day -- because that's where he's going to function, that's where he feels frustrated, that's where he feels good.

He also asked the teachers to validate the QRI scores against their own knowledge of each child. During the individual teacher meeting in which the QRI was discussed, Mike said that his procedure was to get the teacher's overall feedback about the scores and then ask, as he went down the class roster, "Is this true for Johnny?"

Mike's individualized approach to instruction was buttressed by his belief in the importance of individual motivation. Although he did not intervene in teachers' instruction, he strongly encouraged teachers to "find out a way to put two and two together. Find out what Jimmy's carrot is and use that carrot to get him to learn and it'll be great." Student motivation was the primary reason that Mike was supportive of the Accelerated Reader program. He called the program "...fantastic. Unbelievable" and continued, "Kids that stand in line to take a reading test, comprehension test, you know, who's gonna stop that?" Mike supported the teachers wholeheartedly in using it and, in the process, probably contributed to the "sacred cow" status attributed to it by Liz, the district curriculum coordinator.

Mike's belief in the uniqueness of teachers and students carried over to his beliefs in the uniqueness of schools. Like the other elementary schools in the district, Lincoln also participated in the Accelerated Schools model. Mike particularly liked that the

process was flexible, that they could “fit it to our building, our staff, our kids.” He said that Lincoln chose Accelerated Schools primarily because “...it’s a process and not, here comes another curriculum.” The school had already “done some vision work” and begun setting goals the year before they started Accelerated Schools, and so they were also able to use that “instead of starting from ground zero.”

Mike had a strong belief in the individual capability of each of his teachers and the importance of personal motivation. As a result, he was firm in letting each teacher teach in the way he or she thought best, even at the expense of school harmony. Mike’s views, coupled with the teachers’ historical self-reliance and the district’s support for strong school self-governance, created a school context which encouraged the development of teacher autonomy but not necessarily group cohesiveness. Mike’s views of the importance of recognizing the individual in teaching and learning also supported the individual focus of the state second grade reading test.

#### State Assessment Mandates

As a previously low-scoring school, the staff at Lincoln was especially cognizant of the need to improve students skills. Mike commented, “... yeah, the WASL is obviously very important and we’re trying our tricks of the trade to get our kids to do a good job.” Mike described the purpose of staff meetings as getting the EALRs into “everybody’s mind so that we’re just automatically thinking about the EALRs as we make lesson plans and as we talk daily with the kids.” He also described how each teacher in the building watched the fourth graders take the WASL during the spring 1998

administration for at least some portion of the test. The teachers came away from that exercise with the realization that “it’s a reading test, so to speak,” and as a result Mike (and other teachers) felt that teachers were doing more reading and writing in class. Mike also described how state scores were a real “jumping off piece” for the third grade team (classroom teachers, instructional assistants, Title 1/special education teachers, English as a second language teacher) when they received ITBS state scores that were lower than expected. Mike provided the opportunity for the team to get together to brainstorm how to better serve the students and reported, “It was just awesome to watch them work together (as they discussed) how better to help these kids when we’re already doing what we think is all we can do.”

Mike used the school’s increasing test scores as evidence that the teachers were doing a great job. He said,

Our kids are reading and the scores are going up. Slowly, but they’re going up and so you know I just keep pressing, you know we’re reading, we’re writing, we’re writing, we’re reading.

At the same time, Mike acknowledged his ambivalence about the importance of state tests when he said,

Will our kids be prepared for life? And then at the same time I’m flipping back to the other side. Does truly not passing these tests mean that they won’t be successful? And I think that there’s going to be a whole lot of kids that may or may not pass the high school test that will still be successful in life.

Mike saw the state-mandated second grade reading test in the context of the WASL and felt that the test was well-received in that “now we’ve got a better judge if our kids are going to do well on the 4th grade test.” He also acknowledged, however, that there was a feeling among staff that “here comes one more thing from the state. Do we have time for it? Do we teach to the test? All the same questions that I’m sure came up for the 4th grade... Yeah, it’s real pluses and minuses.”

Although Mike was not “that familiar” with the QRI, he saw it as “one more assessment piece ... I think it does a good job of showing legitimate growth or non-growth.” Mike reviewed the QRI, along with the STAR test results and the Gates McGinitie results with his teachers. For Mike, although the test had an individual focus lacking in the other two district tests, the QRI was one test among several. He said, “The QRI is as important to me as any assessment.” He thought that the test was having a positive impact because of the accountability, both for his teachers and for himself.

### Summary

Millersville’s school-oriented governance model was a good match for the principal of Lincoln and his staff. The principal, with a traditionally self-reliant staff, believed strongly in his teachers’ competence. Although he supported his teachers in many ways, he had a “hands off” attitude in regards to their instructional methods. Mike was given leeway by the district climate to support his teachers’ autonomy and to rely on their instructional decisions, even when they ran counter to district expectations.

Lincoln School engaged in many activities in direct response to new state standards and assessments. However, at the same time the principal recognized the limited nature of the assessment measures. The combination of school activities, teacher autonomy, the number of tests given and Mike's moderate attitude towards the mandated state tests created a context in which teachers were expected to use their own judgments to make sense of the required tests, either individually or in grade level teams.

#### Webster – A School in Millersville Aligned with State and District Reforms

Webster School was also under construction during the year that I observed. Large dump trucks beeped as they backed up and the students tracked mud into the classrooms as a new school rose next to the old. Although the outside of the old building looked dusty and worn, inside the wide hallways were brightened with student work and busy with tables and desks where students worked with specialists or instructional assistants. The staff seemed noticeably happy to be there and were unfailingly friendly and helpful to me.

Surrounded by small homes on large lots, the area immediately surrounding the school had an almost rural feel, except for the noise of freeways which could be heard when the wind blew in the right direction. Nonetheless, like Lincoln, this was an urban school serving many immigrant and transient families. The students at Webster spoke 22 different languages and came from 20 different nations. "Minority" students made up 66% of the 353 students, and 79% of the students received free or reduced price lunch. In the year prior to my observations, 99 students withdrew and 135 new students entered

the school during the school year. Despite these challenges, reading scores on the WASL had risen dramatically in the past three years from 32.8 % of the students meeting standard in 1997, to 57.4% in 1998, to an amazing 71.2% in 1999.

Vicki, the principal of Webster, had been in education for 28 years. After many years in Oceanview district as both a classroom teacher in grades K-6 and a Chapter 1 reading and math specialist, followed by a stint as an assistant principal in a neighboring suburban district, Vicki was in her third year as principal of Webster elementary. Holder of a masters degree in curriculum and instruction with a specialization in multicultural curriculum from a nationally-ranked university, she came to her job with clear ideas about reading instruction. She strongly supported state reform, especially the WASL, and many of the Millersville curriculum initiatives. She saw her job as one of helping people to help students and said, "...you can reach a lot more people, a lot more kids, if you're not just within your own classroom. You miss that but you can spread yourself a lot more." For Vicki, the bottom line was kids. She spent a lot of time on the playground and in classrooms talking to children and commented about the principalship, "Not the easiest job but the rewards are the kids, when you see them grow and do well."

### The Influence of State Reforms

Vicki was a strong proponent of state reform. Student work captioned with the appropriate EALRs hung on the walls outside her office, graphically illustrating her support to teachers, parents and students. Vicki "loved" the WASL and said, "I think it's what we should have been doing all along. It is totally different than any other

standardized test. It's true-life things that kids need to be able to do." School money was used from the beginning of Vicki's tenure to help teachers match their instruction to the WASL, and it supported such activities as professional development and release time for teachers to align their curriculum to the EALRs.

One of the teachers reflected Vicki's enthusiasm when she noted that at Webster the WASL was not a fourth grade test, but a school test saying, "Everybody is involved in one way or another and it affects everybody." During the spring of 1999 Webster had a WASL kick-off assembly and K-5 teachers arranged their schedules (with the help of substitutes) so that they could each administer the test to a small group of the 4th graders that they had previously had as students. The plan to do this originated when they met as a building team and with Liz, the district curriculum coordinator, to figure out ways, according to this teacher, that "would help the kids get through the test." Both of the second grade teachers in this study thought that this organization of test administration was an important factor in the students' improved scores. (The increase in reading scores at Webster, for example, far exceeded the increase at the other schools.) One teacher said about this scheme, "And they were kids that I had had in the past so they were already ...bonded to me and they knew I cared about what they were doing so ... that helps raise it (the scores)." This teacher also noted that this experience helped all of the teachers at the school to become aware of what "exactly is on the test. Because ... you can go to a workshop and read through the test yourself, but until you actually see your kids working on those tests do you realize what it is that they really need to do to be successful.".

Vicki was also politically astute about state reform efforts and confident about her own knowledge about reading instruction with regard to the reform. Because of its low 1997 reading scores, Webster received a state K-2 reading grant. This money, legislated at the same time as the second grade reading test, required that schools use the grant money to purchase phonics-based materials only. Vicki noted the political nature of the phonics requirement and, confident in her own knowledge, defined what was important for Webster to do with the money. Rather than buying phonics programs, and in line with her belief in the importance of integrated literacy instruction, Vicki let the teachers buy “just like lots of books because I felt, you know, any book is a phonetic reader.” The teachers found lots of books and games that did explicitly meet the phonics requirement as well but teachers only purchased them “...if they felt it was good and it truly worked for our kids.” Vicki was willing to “stretch” the mandate, in her words, to do what she thought was best for the students.

#### District Programs at Webster

Vicki also supported the district curricula in reading and math, valuing them particularly for their alignment to the higher level skills she saw required by the WASL. Although she “would not have chosen” Scholastic simply “because I just truly believe in integrated curriculum where you don’t have a set of books to read. I think it just should be all integrated,” nonetheless she thought that Scholastic was “pretty good.” She believed that Scholastic was chosen because it had “higher level kinds of activities, projects and thinking skills” as well as non-fiction, which aligned to the WASL. Perhaps

in reference to teachers in the school who were not using the Scholastic materials very much, she added that she thought that teachers needed more professional development in using the Scholastic materials, because, "... there's always so many more ways to use a series than you know you can think of when you first start using it." Vicki also valued the district curriculum SAIL, believing that it led to "higher level thinking skills" and saying that she thought the biggest district influence on the building was the SAIL staff development. She said that it was an expectation that all that teachers would teach SAIL and commented that it was in use "throughout the building."

In Vicki's mind, the second largest district influence was its support of Accelerated Schools. Like the other elementary schools in the district, Webster participated in the Accelerated School program. According to Vicki, her staff chose Accelerated Schools because of a short timeline, because that was the one the superintendent recommended and because she (Vicki) liked it. Vicki liked Accelerated Schools because

I'm one that believes that from the custodian, the cook and any classified staff, the office, anyone here, you make an impact on a child's life and so I liked that program because it had everyone involved in it. I liked it because it didn't say take this program -- it will change your scores. Because we're all different.

Vicki saw the district as strongly supporting the project in her school, both in terms of validating the program's importance and in terms of providing financial assistance to make it happen.

Vicki described Accelerated Schools as first helping the staff to deal with governance issues -- they had to learn how to make their building run effectively before they could work on instructional issues. She thought that Accelerated Schools had helped the school learn to work together as a team and make meetings productive. She felt that Accelerated Schools mirrored her own vision of instruction in that rather than laying out an instructional format for teachers to follow, it asked staff to "discover just like I want the kids to do, what's going to work best for you in your building with who you are." Vicki also thought that the process had an impact in terms of the way staff would ask, when considering a change in their instruction, "Well, what does the research say?" because of the research process in the subject area cadres (literacy, math and science) to determine the strengths, problems and possible solutions for these subjects at the school. During the 1999-2000 school year the literacy cadre decided to focus on a writing action plan, leaving reading for the following year.

The Accelerated Schools model also created a problem for Vicki with the multi-aged Grade 2/3 team. One of the principles of Accelerated Schools stated the importance of using the strengths of the people in the building. She said she had been pleased when the Grade 2/3 teachers told her that they were using something from Accelerated Schools, "teaching to our strengths," and thought it was fine that they were each going to be teaching one subject matter area (math, literacy, social studies/science). She neglected to ask enough questions to discover that students were ability grouped and moved with their ability level across all three classes. Vicki strongly disagreed with this approach and said, "It's totally against the philosophy of multi-age. Totally against good research."

However, by the time Vicki realized what was happening, the program was already running and she said that to ask the teachers to change would not have been good for the students or staff morale. At the end of the 1999-2000 school year, Vicki broke up the team. The teachers told me that they were assigned to different grades and to rooms in the new building that were separate from each other. Vicki also asked them to mentor less-experienced teachers during the coming year.

Vicki was also in conflict with the Grade 2/3 team over Accelerated Reader (AR). AR was a key aspect of the literacy program for the team, and one of these teachers was primarily responsible for its organization in the school. Like Liz, the district curriculum coordinator, Vicki had some concerns about AR saying,

There's some research coming out about the AR too, some negative things...

Getting into the research mode, we need to start looking at some of that research too.... I need to make sure I keep a check on that. AR is strictly to help children to learn to want to read and that you can read for many purposes and to give them practice reading. It is not a reading program a such.

At the time of these observations, however, this conversation had not yet begun.

Vicki strongly supported the state reforms and used them as a way to judge the appropriateness of various programs. She believed that the district curricula and staff development program were generally good because they were chosen with the WASL in mind, and she expected her teachers to use them. In the case of the Scholastic materials, which were not frequently used by some teachers and which she actually didn't support whole-heartedly, she attributed the lack of use to a need for more professional

development, not a defect in the materials. Vicki's confidence in her own knowledge gave her the background to question some of the choices that the Grade 2/3 teachers made, such as Accelerated Reader and ability grouping for instruction across classes, and put her in conflict with the team. In her support of state reform, district curricula and questions about Accelerated Reader, Vicki's stance was aligned with Liz's in the district office. Vicki's clear curriculum stance also situated the QRI within a school context of strong instructional expectations.

#### The QRI at Webster

According to a long-time teacher, second grade had been an emphasis area at Webster for many years. The school used to have a volunteer program, HOST, in which the reading teacher recruited volunteers to come in four days a week to work with primarily second grade students. The teacher added "It's just ... always been the philosophy here of our reading teacher and our curriculum directors that (in) first grade they're still kind of feeling their way ... but by the end of second grade they need to be pretty competent readers."

As a result of this philosophy, and the fact Vicki had a background in reading instruction, I expected that she would strongly support the QRI. However, she did not. She thought that the purpose of the test from the state perspective was "really more about numbers and let's see if people are improving, where are they, how are our kids doing in terms of what these tests say." Seeing it primarily in terms of an unfunded state mandate, she did not use it at the school level in conferences with teachers or to determine resource

allocation, and she did not think that teachers liked it, stating sarcastically, “They’re *overjoyed* that they have it.” Later she added, “It’s one more thing for them to do. But since you have to give it, it’s at least another piece of data.” Vicki provided subs for the Grade 2/3 teachers to give the test because it was “one more stress. I didn’t think was worth it.” Vicki also alluded to a notion of the QRI’s unreliability in relationship to other tests when, referring to the fact that the kids were tested individually she noted “Kids know it’s important so therefore they probably do a little better because they’re sitting with their teacher one-on-one.” During the interview Vicki mentioned that the QRI was just “another piece of data” five times and, given that she did not make a distinction between information provided by this assessment and the other assessments that were given by the Grade 2/3 team (the Buck County Test – a locally-developed reading test, the Gates McGinitie, the STAR test and all of the individual reading reports from the Accelerated Reader program), her statement is understandable. Vicki concluded this section of the interview saying; “I don’t see this great impact other than on teachers’ time, probably. But not in terms of, would it change a child’s score totally around for the year because they took the QRI with the teacher.” Vicki saw the QRI in terms of its utility for raising student scores, not in terms of its utility for providing useful information for teachers.

Unlike the other state reform mandates, Vicki did not support the QRI. She thought that it took up teachers’ time, added stress, took money from her budget, and that the main purpose of the test was to provide numbers at the state level. The information provided by the QRI was also suspect because, unlike other tests, it was administered

one-on-one. In the end, the principal did not use the QRI scores for teacher conferencing or school level decision-making.

### Summary

The principal of Webster strongly supported state reform efforts as well as most district curriculum initiatives. She expected teachers to follow district curriculum guidelines and supported school-wide activities that would increase teachers' understandings of the EALRs and the WASL. At Webster School the QRI was placed in a context of clear instructional expectations. However, despite the school's history of an emphasis on second-grade student reading achievement, and despite her support of state reform and the WASL, the principal of Webster was not supportive of the QRI. Her understanding of the QRI as a stress for her teachers and just "another piece of data," when the teachers already collected so much data, reflected her belief that the purpose of the QRI was primarily to collect scores for the state. She did not understand that the QRI could provide useful instructional information for her teachers. Even though the district curriculum coordinator understood and supported the potential that the QRI offered to develop classroom-based assessments, the principal of Webster did not. The principal's conflict with the second/third-grade team over instructional issues also created a climate which made the use of second-grade QRI data at the school level, either in terms of conferencing with teachers or using the data to help make school-wide decisions, less likely.

Millersville District, Lincoln and Webster Schools – Small and Independent

The small scale of Millersville gave the district little capacity to respond to state reforms. With a history of school independence, the district opted to enhance the schools' decision-making capacities with a whole school reform model, Accelerated Schools. The district's limited capacity, combined with the enhanced decision-making capacity of the schools, made it difficult for the district curriculum coordinator to influence the use of district-wide curricula and assessments. This was especially true in the case of the QRI, which was administered by second-grade teachers in addition to the Gates McGinitie and STAR test, assessments that had previously been established by the district. Although the QRI was recognized by the district curriculum coordinator as a way to help teachers learn about classroom-based reading assessment, the district did not provide professional development about this potential use of the QRI. Lacking instruction or district pressure, principals recognized it as simply one assessment among many.

Millersville had two district-wide literacy curricula in place and one long-term literacy professional development effort that involved every elementary teacher in the district. These district efforts provided a context for instruction and assessment within which the QRI was situated. SAIL, with an emphasis on six reading strategies, encouraged teachers to interpret students' reading behavior using the SAIL strategies. Accelerated Reader, which used multiple-choice questions to measure literal reading comprehension, promoted the view that questions provided sufficient evidence of

students' comprehension. SAIL and AR had the potential to impact how teachers interpreted and used the QRI.

Although both schools participated in the Accelerated Schools model, the principals at the two schools held quite different approaches to governance and curriculum. In one school the principal considered the teachers instructional experts and felt it was his job to support them, even if what they wanted went against the district direction. This principal took a practical but low-key stance towards state testing, and showed that he valued the QRI results by using them to conference with his teachers about individual children. At the second school, the principal had clear ideas about literacy instruction which aligned with the WASL and district curriculum initiatives, but which was, in some instances, in conflict with the instructional approaches of the Grade 2/3 team. In this school the principal did not use the QRI results because she saw them as suspect numbers that did not add any further information to the data that had already been collected.

Although the QRI filled a perceived need on the part of the district curriculum coordinator, this perception was not communicated to the schools. As a result, the QRI did not play a prominent role in principals' decision-making in Millersville.

### Oceanview and Millersville Districts -- A Cross-Case Comparison

#### The Districts

Both Oceanview and Millersville were urban districts with diverse populations. They both had school-centered decision-making processes in place to various degrees.

The two districts, however, responded in different ways to the state second grade reading assessment mandate. For Oceanview, a large district with a large central office staff, the state second grade reading test filled an accountability need. With many different curricula in place in the various schools in the district, the DRA fit into the developing district standards and assessment system and provided a measure of consistency and accountability across a large varied district. The district supported the DRA wholeheartedly by expanding the test to other grades and by providing teachers with many forms and directives about the test. Although the Oceanview district reading specialists saw the usefulness of the test in terms of providing a model for classroom-based assessment and the development of teacher knowledge about reading, this message was easily lost in the district's first and on-going message about accountability.

Oceanview schools also used the DRA for teacher accountability and as a way to distribute resources. As a result, teachers generally perceived the test to be high stakes.

Millersville, on the other hand, as a district with a very small central office staff, had already met the need for second-grade accountability through the STAR and Gates McGinitie tests. With little district capacity to inform teachers and administrators about the test, and no district pressure, the QRI was not used to distribute resources in the schools and was used only minimally for teacher accountability in one school. Although there was diversity among Millersville teachers in terms of the kinds of curricula they used (basal readers, sets of trade books), there was also consistency in district curriculum and instruction (SAIL staff development and Accelerated Reader). Even though the district curriculum specialist recognized the potential of the second-grade test to support

teachers in developing classroom-based assessment, she lacked the resources and the power to support it. The test was considered low-stakes in the district by teachers and administrators and was not used to distribute resources or evaluate teachers, although it did serve as a focus of a teacher/principal conference in one school.

In both districts, the state second-grade reading assessment was one of a number of reading tests recently required by the districts at the second-grade level. In Oceanview these tests were all district-developed classroom-based assessments (a phonics test, a sight word test, and a WASL-like comprehension test) and the results did not have to be submitted to the district. In Millersville the tests were commercial (the Gates McGinitie and the STAR) and the results were sent to the district office. In Oceanview the presence of the other required tests did not detract from the importance of the DRA, which was the only reading test that had to be submitted to the district office, while in Millersville, with all of the assessments submitted to the district, the QRI assumed no special importance. There is a Catch 22 here. On the one hand, district accountability, as present in Oceanview and lacking in Millersville, seemed to be a necessary factor in creating a sense of importance about the test. On the other hand, Oceanview's accountability messages interfered with the message about the usefulness of the test to inform teachers' instruction.

There were also similarities in the ways that the districts dealt with the tests. Both districts provided inservices about the second grade test. The focus of the inservices was on how to administer the test rather than how to understand or use the test results. As a result, teachers were left to develop their own interpretations based on their previous

knowledge. In essence, the districts provided inservice on the topical aspects of the test - - what it was and how to give it -- but neglected the more substantive issues surrounding the test -- how to interpret and use it -- which were required to help teachers become more knowledgeable about reading and assessment (Spillane & Thompson, 1997).

In sum, it appears that there were two major factors that influenced the districts' responses to the state second grade test. The first was the degree to which the district used the test for accountability purposes. In Oceanview, the DRA fulfilled this need; in Millersville it did not. Without the pressure caused by district accountability, Millersville administrators did not take the test very seriously. Secondly, the capacity of the district to support the second grade test also played a role. It is relatively easy to mandate tests and more difficult to provide the resources necessary to change teachers' practice. Oceanview had the financial and personnel resources to support the accountability aspects (the bubble sheets, the test directives) of the DRA across the district, but lacked the resources to fulfill the other purpose of the test -- the development of teacher knowledge about reading and classroom based assessment. At the same time, Millersville had the financial capital to develop teachers' understandings about the test but, because of other on-going district initiatives, lacked the time and human resources to pull it off.

### The Schools

Spillane and Jennings (1997) note that it is useful to think of policy implementation as an "opportunity to learn" and indeed, one of the intentions of the state

second-grade reading policy was to educate teachers about the reading process and about instruction. The districts provided “opportunities to learn” by offering the second grade reading assessment inservices. The schools also provided various “opportunities to learn” such as school-level professional development (Valley View) and interaction with the principal or colleagues around the test (Valley View, Oak Park, Lincoln).

Collegial interaction provides teachers the opportunity to test and change their understandings (Spillane and Thompson, 1997). Valley View was purposely attempting to move from a “congenial atmosphere” to a “collegial atmosphere” through school-based staff development in literacy that was congruent with the DRA. This provided a rich context for the development of teacher knowledge about reading based on the DRA. In contrast, Oak Park, the second Oceanview school in the study, had a contentious faculty and sent a school team to the district literacy inservices. The faculty was not collaborative, at least at the second grade level, and the district inservices, with a focus on instructional strategies, did not provide a context to help teachers make more sense of the DRA.

In Millersville, the Accelerated Schools model helped schools to become more collegial through the development of school norms for discussion as well as through specific subject matter cadres. Although literacy cadres were part of the process in both schools, the QRI was not a part of the discussion in either of them. In one school the focus of the literacy cadre was primarily on writing; in the other school the focus of the literacy cadre was on the revision of existing instructional materials.

The state second grade reading policy was intended to serve as a way to get resources to children who were struggling readers. It was also intended as a way to educate teachers about the reading process and instruction. However, state policies are interpreted by districts and schools in light of their own needs and understandings. As a result the outcomes of any one policy are often quite variable. This was the case with this policy. In one district that attached district accountability to the test, schools allocated resources for children on the basis of it. On the other hand, in this same district messages about accountability created the perception of the test as high-stakes and masked the test's purpose of educating teachers. In the other district, which did not attach accountability to the test, resources were not distributed on the basis of the test and the state assessment was perceived as one test among many. The test's educative function in both districts was also limited by the lack of state funding to support district inservice or release time for the test while the view of it as an "unfunded mandate" negatively impacted administrators' perceptions. Schools, in terms of their staff development programs or degree of collegiality, also impacted the efficacy of the policy.

The degree to which the second-grade reading assessment provided opportunities for teachers to learn was influenced by state, district and school factors. What teachers learned, and what affected their readiness to learn, is the other side of the equation and the subject of the next chapter.

## Chapter 5

### The Teachers

The state second-grade reading test gave teachers an opportunity to learn about their students as readers, and an opportunity to reflect on their own reading instruction and assessment. This chapter discusses what teachers learned from the second-grade reading test and how they used what they learned. It also discusses the ways in which teachers modified the test based on their beliefs and knowledge about reading and assessment. The influence of district and school contexts on teachers' interpretations of the test is examined, as are the reasons why teachers generally valued the test. Underlying all of the decisions teachers made were assumptions about the role of assessment in instruction.

#### Learning From and Using the Second-Grade Test

Teachers learned from the second-grade reading test when they expected it to deepen their existing understandings of children or when it provided a way to locate students on a continuum of reading proficiency. Teachers also learned when they experienced dissonance between their expectations for students and how students actually performed on the test. Dissonance was provoked by conflict between student results on previous assessments and the test, or when students performed well on some dimensions of the test and poorly on others. Dissonance forced teachers to acknowledge the complexity of the reading process. One teacher, who did not find the test useful and who

did not experience dissonance, did not learn from the test. In all cases when teachers learned something, it impacted students' instruction.

### Deepening Understanding

The structure of the test – oral reading and retelling, with scoring based on several dimensions – gave most of the teachers an opportunity to learn more about their students than they had through the assessments already in place in their classrooms. For four teachers the test either replaced a similar informal reading inventory that they had previously used, or added to the individual assessments that they were already administering. For these teachers the test also was a good match with their instructional programs. These teachers saw the test as a way to deepen their knowledge of the individual reading abilities of their students. Jean, a teacher at Lincoln School in Millersville district, provided an example of a teacher who fit this situation.

Jean replaced her usual Informal Reading Inventory with the Qualitative Reading Inventory, the second-grade test chosen by her district. With a background in Early Childhood Education, she had a strong belief in the importance of individual instruction and assessment. As a classroom teacher she used a battery of commercial tests, but said that she “mostly” relied on individual observation of students. She said she liked the QRI because she could use it to show growth in students' accuracy and rate and because it gave her a chance to “Ohh and Ahh” with the students over their improvement, giving her a way to “touch each child.” Later she added, “When I do the QRI it's just that child and myself and I can truly see what they're struggling with. I can get the STAR (a multiple-

choice computer-based comprehension test, cite) results, but I still don't know what they're struggling with."

The QRI gave Jean the opportunity as well as a structure within which to observe particular reading behaviors. For example, speaking of her student, Jabar, Jean said that she "already knew he had a lot of gaps in his learning" but that the QRI gave her more specific information, such as needing to work more on sight words and "ed" endings. Jean was also surprised to learn, as a result of the QRI, that two of her top readers were less accurate than she had expected. As a result, Jean determined that she would focus more on accuracy with these readers.

Jean relied heavily on the progression in the basal series to drive her instruction. The QRI was a good match with the goals and format of the basal reader she used in that it was organized by grade level and there was a focus on word identification skills, fluency, and comprehension. The QRI provided Jean with information that impacted her interactions with particular students but did not change the over-all course of her instruction.

Like the other three teachers who saw the test as a way to deepen their knowledge of their students' reading abilities, the second-grade test aligned with Jean's beliefs and practice around literacy instruction and assessment. Particularly important were these teachers' beliefs in the importance of individual assessment. Because of this alignment, and the structure of the second-grade reading test around four dimensions of competency, these teachers were able to add specificity to their previous understandings

of students. These teachers also used the test at the end of the year to share with their students the reading progress they had made.

#### A “Locator Device”

All of the teachers in the study saw the second grade test as a “locator device” (Shepard, 2000a), a way to place students on a continuum of reading ability. Two of the teachers used the second-grade test mainly as a method for determining at what level students were reading so they could be assigned to an appropriate reading group. Several of the teachers used the assessment to help them validate appropriate grade-level instructional text levels for students. Other teachers saw the second-grade test as a “locator device” but did not learn anything new from the information it provided because they already had plenty of other assessments in place to meet this need.

The focus on the test as a “locator device” reflected a “sorting and classification model of ability testing” in which test results are used to “assign students to gross instructional categories rather than having the test tell something particular about what a student knows or how he is thinking about a problem” (Shepard, 2000a, pps.8-9). This attitude was held, at least in part, by all of the teachers and had ramifications in terms of how teachers thought about and used the assessment.

In the first place, by emphasizing that the role of testing is to place students on a reading continuum, the assumption that reading development proceeds along a one-dimensional continuum is emphasized, downplaying the complexity of the reading process. Second, this perspective values the second-grade test primarily as a means for

evaluating the performance of students, as opposed to being a way to learn more about students as readers. All of the teachers held these views, while at the same time and to varying degrees, holding views about other purposes for the assessment.

When teachers used the second-grade test as a way to determine placement of students in reading groups, they sometimes expected that the assessment would supply an achievement level and not necessarily new understandings about their students (Shepard, 2000a). For example Dolores, a teacher at Oak Park School in Oceanview district, grouped for instruction. She referred to her use of the second-grade assessment as “diagnostic” because she was going to use it to assign students to reading groups by their reading level as determined by the DRA. Shirley, another teacher at Oak Park who used the assessment to group for instruction noted in this exchange:

Interviewer: How did you use the results?

Shirley: This year to place students.

Interviewer: To place students. Do you use them in any other way besides placing students?

Shirley: You know, I don't really use it any – just to know where -- basically just to provide the kids with the necessary materials for the level they're at.

An emphasis on accountability may also have focused teachers more on performance levels than on understanding the complexity of the reading process. Jean, who used the QRI results in combination with other tests (the Gates MacGinitie and the STAR test) to report to the school board, put reading levels as reported by each test on a spread sheet, so that she could report students' progress. In Oceanview, the district's

emphasis on accountability may also have focused the teachers on their students' performance levels at the expense of recognizing the complexity of the reading process.

Shirley, for example, noted that the DRA was helpful to her saying,

Now we know what we're working towards, and which I think is helpful to me to make sure that these kids are at that level. And again, this puts a little more pressure on us too, because making sure those kids that aren't quite at level then we have to really put a lot of work into making sure they're at least close to that.

All of the Oceanview teachers felt district pressure to get their students to the second-grade standard. They all referred to their students at times by the levels at which they were reading saying, for example, "Oh, he's an 8," which referenced a short-hand, global way of thinking about their students' reading. This undifferentiated type of knowledge about students' reading was representative of the type of information teachers were used to receiving from other reading tests.

Four of the teachers in this study learned from the second-grade test because they expected it to deepen their understandings about children. For these teachers there was congruence between the second-grade test and their beliefs about literacy and assessment, particularly in terms of their beliefs about the importance of individual assessment. In addition, the structure of the test focused teachers on the dimensions of reading, which helped them develop more specificity in their understandings of their students as readers. Although all of the teachers used the test as a "locator device," when this was seen as the primary purpose of the test, as it was for those teachers who used it to group for reading instruction, teachers framed the results in terms of achievement levels as opposed to the

multi-dimensional process that the test was intended to represent. A district emphasis on accountability also seemed to focus teachers' attention on the performance aspects of the test.

### Dissonance – Between Previous Assessments and The Second-Grade Test

The teachers discussed above all expected to learn something from the test. Dissonance, however, was a second factor in teacher learning and occurred whether or not teachers expected to learn from the test. One way that dissonance occurred was when there was conflicting information between teachers' previous assessments and the second grade test. This type of dissonance was particularly noticeable for Gloria, who relied heavily on the Accelerated Reader (AR) for her information about students.

Gloria, who had been teaching for 30 years, taught reading to all of the second and third graders in Webster School in Millersville. She did not expect to learn much from the QRI because she had already tested her students on the Buck County test and the STAR test. As a result of these tests (and their math scores) the students had been placed in reading groups and were participating in the Accelerated Reader program.

Of all the assessments she used, the Accelerated Reader program had the most influence on Gloria's thinking as she predicted students' performance on the test. For example, before we listened to Albert read, Gloria noted that he had gained 10 points in the Accelerated Reader program during the first two weeks of school. This meant to her that he was reading avidly, having read several books and passed the accompanying Accelerated Reader tests in a short period of time. In addition, Gloria noted that Albert

was working on “two point” books, which meant that he was reading higher-level books. She said, “I don’t feel it’s necessary to give the test (the QRI) to Albert because I already know what he’s doing.”

Gloria, however, was surprised at how poorly Albert did on the retelling, a section of the QRI. In reflecting on Albert’s performance, Gloria noted that the retellings are “the part that most surprises us sometimes” and, indeed, Gloria and her teaching partner Sarah were surprised by the retellings of two other children, who retold the passages better than they expected. Gloria credited Albert’s poor performance to the fact that “he really didn’t know what to expect or what I was looking for when I was asking him for the retellings.” She felt that he would have done better if they had used the comprehension questions, which weren’t so “open-ended.” Gloria felt that Albert would improve in this area because she had already planned that his reading group would do a lot of work on retellings. Nonetheless, as a result of the QRI, Gloria planned to ask Albert’s mother to have him retell chapters in books that he read at home.

In another instance, Gloria used the QRI in combination with the Accelerated Reader to make a specific instructional plan for a child. David was a fluent and accurate reader (except that he skipped three lines at various points in the text without recognizing he had done so!) who did an incomplete retelling. His scores put him in the “borderline” range. (A borderline student according to the state rubric was one who was not substantially below grade level, but one whose progress should be watched.) At first, Gloria decided that this child didn’t need extra help “because of the kinds of errors he made and had he been given a marker he would have done all right. And the retelling

I'm not worried about either because, as I said before, he'll receive instruction on that." However, later in the morning Gloria remarked to me that David was not passing his Accelerated Reader tests either and so she had decided she would give him extra help by having an Instructional Aide work with him one-on-one to help him "learn to read for meaning." (In this scheme, David read two sentences and then retold what he had just read.) Later in the year, Gloria reported that it took several months before David understood how to retell and before he started passing the Accelerated Reader tests. In this case, the retelling section of the second-grade test enabled Gloria to reflect on David's comprehension and, coupled with the Accelerated Reader results changed her instruction.

All of the commercial assessments that Gloria used relied on multiple-choice answers to literal comprehension questions. These test results gave teachers a grade level designation for reading, but they did not measure students' ability to construct meaningful recounts of a passage, as required by the retelling exercise for the second grade assessment. Although Gloria felt that the second grade test provided mostly "verification" for what she already knew about children, the individual nature of the test coupled with its use of retellings caused Gloria to change her understanding, and in some cases her instruction, of students' comprehension. It also highlighted the differential performance of students in situations when different comprehension assessment measures were used.

### Dissonance Caused by Conflict Between Dimensions

Another source of dissonance for some of the teachers was based in the dimensions used for scoring the test. When a student scored high in some dimensions and low in others it created a conflict for teachers, especially if they had previously judged the student proficient based on only one or two dimensions. The dimensions of the test illustrated the complexity of the reading process.

Louise provided an example of a teacher who was surprised by a disjuncture between her previous observations of a student as well as a disjuncture between the dimensions on the second-grade test. Louise, a first year teacher at Valley View in Oceanview district, initially started reading with her student Howard at a low level in the DRA because “when we talk about things (in class) he’s like whose (sign for whizzing by his head) -- he’s in another world. So that’s why I really thought he was (Level) 8.” When Louise tested him, he read Level 8 so easily that she immediately moved him to the grade level passage (Level 18). He also read this quickly, accurately and with expression. However, Howard could retell very little of the story, even with her probing questions. At the end of his session she remarked to me:

The retelling. That was telling. So the kid blew through this -- probably the best out of anyone, and could not tell me what was going on.... He could read but he could not retell the story. He has no idea what’s going on. Even looking at the pictures. He couldn’t tell me. And that says a lot.

This observation started Louise on the road to beginning to solve the puzzle of Howard. He initially received one-on-one attention from the ESL aide in her classroom to help

with his comprehension and eventually became a focus of concern in terms of possible autism.

As a first year teacher, Louise may have especially benefited from giving the second-grade reading test. She had learned how to administer running records in her teacher education program, but had not done this in her classroom, and so the format of the DRA gave her a reason to implement what she already knew. Eager for information about her students, the one-on-one format of the test gave her an opportunity to step away from her class and notice things about each of her students that surprised her and that she had not noticed before in the group setting.

Gloria and Louise provided two examples of teachers who learned from the second-grade test because of the dissonance it provided either between their previous assessments of students, or because there was an unexpected differential in student performance in the dimensions of the test. If teachers didn't expect to learn from the test, or if there wasn't some kind of dissonance between the test and teachers' expectations, they didn't learn from it. Such was the case of Jody.

### Eliminating Dissonance

There was one teacher in the study who reported during the think-aloud that she did not learn anything new about her students from giving the second-grade reading test. According to Jody, who had been teaching for 12 years at Lincoln School in Millersville district, "the test (the QRI) does not add a whole lot of information to what myself, or

Marian (instructional aide) haven't picked up before we administer it... It's just one more piece of paper. No surprises."

There were several possible explanations for Jody's stance towards the test. First, Jody administered many assessments to her students. Prior to administering the QRI, Jody had listened to each child read individually to "see where their level is" and "checked out their writing to see what skills they have holes in." Listening to students read aloud was also a way for her to focus on her students' use of phonics. Jody also asked the students' previous teacher to rank them. Each child had also taken a Harcourt Brace Jovanovich (HBJ) Power Test (a basal test) which was a vocabulary and comprehension test and which confirmed for her the appropriate basal reader level for each student. They also took the Buck County test, which Jody thought was required by the district (but actually wasn't) and the STAR test, which gave a reading level. Each student had already been placed in a reading group and was reading in the Accelerated Reader program. Jody said about her use of so many assessments,

I take them all into account. I take them all as a piece. And their daily work, and their HBJ Power Tests, which are vocabulary and comprehension, I take that into account too. I don't take any one, because there's so much room for ... depending on the day they took the test, it could be a fluke, but if you take them all you get a real good picture. Along with their daily assessments.

Because Jody gave so many assessments herself, including listening to students read individually, she did not feel that the QRI added anything to her understanding.

Second, Jody's instruction matched most of the assessment. The QRI is divided into grade levels, and Jody thought about her three reading groups at the beginning of the

year in terms of first, second and third grade, designations that matched the QRI. The focus of the QRI on word identification, decoding and comprehension, was also a good fit with Jody's skills-oriented instruction.

Third, Jody did not experience any dissonance because she changed the requirements of the state reading test to match her instructional practice. Comprehension in Jody's classroom was mostly focused on vocabulary meaning, literal understanding and personal response. Instead of using the retelling as a measure of comprehension, as mandated by the state reading test, Jody decided to use the comprehension questions provided with the QRI. She said, "I know who's comprehending and who's not, so it's not necessary and besides, (pointing at the retelling scoring sheet in the QRI with some irritation) this is too cumbersome." (She did, however, also have the state rubric in hand which would have shown her another way to score student retellings.) Jody did not understand or value the retellings, which would have given her an indication of the ways in which her students constructed the meanings of the passages, and she did not think that retellings would help her practice, so she didn't use them. She complied with the test, but didn't use the potential that the test provided to further her knowledge of the reading process.

It is instructive to compare Jody with the three other Millersville teachers. Jean, who was a teacher at Jody's school, also substituted the QRI comprehension questions for the retelling. She said she "hated" the retellings because "it's time consuming and it's extremely difficult to keep up, and I just feel like there are so many other things that are beneficial and it's so difficult." Although Jean did observational assessments of her

students, she substituted the QRI for her usually administered Informal Reading Inventories, and as a result she was surprised at some aspects of each child's reading (although she was not surprised by her students' comprehension). Jody and Jean, who both taught at Lincoln School, also felt that they had enough leeway from their principal and the district to depart from the state guidelines.

Gloria and Sarah who taught at Webster, the other Millersville district school, both used the retelling section of the QRI. They learned the most about their students from the retelling section of the QRI because the results contradicted what they had expected, based on the students' work in the Accelerated Reader program.

The QRI did not provide any new information to Jody about her students because of the number of assessments that she already administered to students, because of the match between her instruction and the QRI, and because she changed the state second grade reading test to match her practice and her understanding of reading. Because she didn't expect to learn anything from the test, and because she didn't experience any dissonance between what she already knew about students and the QRI test results, Jody did not learn anything about her students from the test.

Teachers learned from the second-grade reading test because it gave them the information that they expected. For these teachers, the test met an instructional need. Teachers also learned from the test, however, when they experienced dissonance between their expectations for students and how they actually performed on the test. In all cases in which teachers learned from the test, students' instruction was affected.

### Modifying the Test: The Influence of Teachers' Knowledge and Beliefs

Teachers' knowledge and beliefs about literacy and assessment were reflected in what they learned from the second-grade test. However, teachers also modified the second-grade test based on their beliefs and knowledge about literacy and assessment. As was evident with Jody and Jean, who replaced the retelling component of the test with comprehension questions, this impacted what teachers learned from the test.

### Modifications to the Administration of the Test

Because of the weak professional development surrounding the second-grade test, teachers had to rely on their own understandings of assessment to administer the test. This was particularly evident in the case of Dolores, a veteran teacher at Valley View who served on the Oceanview district committee that selected the DRA.

Dolores received 2 hours of staff development on administration of the DRA. She said, "We got the stuff, the issue was learning to do it. And it, it's just like anything else. You get a new car you have to figure out how all the buttons work. It takes a certain amount of time."

Dolores administered the second-grade test in the same straightforward way that she talked about the staff development. The protocol she used for giving the tests ("When I get ready, when I say go, you can start. OK – GO!") was reminiscent of the group tests that she preferred to give. Dolores' protocol was quite different from state guidelines which said that teachers should talk to the students about what they would be

doing (timing, retelling) and introduce each passage by reading the title and giving a short one-sentence over-view. My experience in administering the test, as well as my observations of teachers, showed that the state guidelines for administering the DRA were more helpful to students than the protocol Dolores used, in that the state guidelines prepared students to remember and retell the text and often gave them some key words in the story.

The DRA also represented an approach to reading assessment that Dolores did not understand. On many standardized reading tests, and in many basal readers, passage complexity is developed by increasing the difficulty of decodable and sight words. On the DRA, passage complexity is determined by decodable and sight words, but also by the match between pictures and text and such language factors as repeating sentence patterns. Children are expected to use more sources of information to understand text than just decoding and sight words, especially at the beginning levels. At the lowest levels of the DRA, teachers are required to read the first sentence aloud to students so that they can hear the sentence pattern and use it as an important cue for reading. Dolores did not read the initial sentences to students at the lowest levels and as a result they were unable to use the patterned language that served as the foundation of the text as a cue for reading the words. Students subsequently obtained lower scores than were warranted.

Dolores' lack of understanding and experience with this type of test showed up in other areas as well. She had a difficult time keeping track of student errors when the students read aloud, and as result missed many errors that the children made. She also misinterpreted one student's errors when he repeatedly skipped words in a patterned

sentence, believing that the student didn't have one-to-one correspondence (the ability to read individual words in text) when it was my observation that, after struggling to identify an unknown word, he was applying her directive to "skip it and go on" to the entire patterned passage.

Dolores was not the only teacher in this study who had misconceptions about the DRA. Neither Shirley nor Louise, relatively new teachers who taught at Oak Park and Valley View respectively, introduced their students to the DRA booklets by giving students the title and over-view statements. Shirley also believed, following her previous understandings of how tests were constructed, that the level designations for the DRA referred to grade level designation. For example, although Level 18 was the standard for second grade, Shirley believed that Level 18 represented the first grade eighth month. In reality, there was no relationship between the level numbers and grade/month designations.

Sally, a 12-year veteran who taught at Valley View, was the only Oceanview teacher in my study who administered the DRA according to the state guidelines. However, Sally's individualized approach to reading instruction was very congruent with the DRA. She understood the ways that the DRA levels were constructed using a variety of different cueing systems, as illustrated in the ways that the books were leveled in her classroom library. However, Sally did not attend the district inservice. She was taught how to administer the test by her teaching partner Betty, who was not in the study, as was the newest teacher in the study, Louise. Dolores and Shirley both used basal readers, with a phonics and sight word base, as the core of their instruction, while Louise, who

had just begun teaching, had many other things to think about the first three weeks of school.

The DRA represented a new type of reading assessment, one that was unfamiliar to both veteran and beginning teachers in Oceanview. Because there was little professional development accompanying the assessment, teachers had to rely on their previous understandings about literacy and assessment to administer and interpret the test. For those teachers like Sally whose instructional programs were congruent with the DRA, this presented no difficulty. But for other teachers whose instruction was built around basal readers and who did not understand the premises on which the DRA was constructed, reliance on previous understandings was problematic.

In some ways, the district's request that teachers simply return the raw data to the central office on a scan sheet, with the district computing the final rubric score, illustrated the district position that teachers simply needed to administer the test, not understand it. However, teachers' beliefs about assessment were so strong that even though they were told how to administer the test correctly during the district inservice, and although the directions for administering it by giving the title and over-view statement were included in the district directions, without an understanding of how the test was constructed, or what it was intended to assess and why, teachers' own beliefs over-rode this information.

#### Modifications to Scoring the Test

Differences between student performances on various dimensions of the test created dissonance for some teachers. However, when teachers scored the tests based on

their understandings of what counted as evidence of reading ability as opposed to following the state rubrics, this dissonance was diminished. For example, Jean, a veteran teacher at Lincoln School in Millersville, scored all of her students after the test, using her own judgment, as opposed to following the scoring rubric. In doing this it was likely that she reinforced her previous understandings of students, as opposed to learning something new that would contradict her expectations. In another example, Sally, who had been teaching for eight years at Valley View in Oceanview district, combined retellings with comprehension questions and students' phrasing to get a composite comprehension score. This hindered Sally in recognizing the different kinds of information that retellings and questions could provide, and in discriminating between students' comprehension and their phrasing.

Shirley, a third-year teacher at Oak Park in Oceanview district, also used student's phrasing as a key indicator of students' reading ability. Annette, a student reading below the second-grade standard at Level 12, gave an inadequate retelling. After previously telling me that "she could not go on" if a student did not comprehend the passage, Shirley moved Annette to the next level despite her poor retelling because of her strong expression when reading the passage. Shirley said, "... she was missing some things but I knew she understood so ... I decided her phrasing was really well so I went on." This was a good call on Shirley's part because Annette subsequently moved up two more levels to a Level 16. At this point she still read rapidly (at 67 words per minute) and she was still quite accurate (97% accuracy) but her phrasing was not very smooth. (She had 13 self-corrections or repeats.) Shirley scored Annette just under standard (2) for phrasing and

just above standard (3) for retelling. Although this student easily met the requirements for passing this level, her lack of phrasing stopped her in Shirley's mind. Shirley explained, "This section was kind of a little slower but she was still getting most of the words and comprehension as well. But the reason I stopped her is because it wasn't quite as fluent as the other."

Shirley used phrasing as the key indicator of a child's competency, and her scoring reflected this bias. As a result the second-grade test did not challenge her belief that smooth reading was the key determinant of the quality of a student's reading ability, and Annette remained in the reading group composed of students who were reading below standard.

#### Modifications Based on Teachers' Purposes for the Test

For Sally and Sarah the second-grade reading test provided an instructional opportunity. They both made modifications to the test that reflected their beliefs about the way that this assessment could be used instructionally with children.

Sally, an eight-year veteran teacher at Valley View in Oceanview district, already had on-going classroom-based assessments similar to the DRA, the test selected by her district, in place in her classroom. Because she already had most of the evidence supplied by the second-grade test, Sally expanded the state second-grade test to include literacy conversations with her students. She asked the students questions about such things as their preferences as readers, their strategies when they came to unknown words, and whom they knew who was a good reader. She also responded to students with her

own observations or experiences. For example, when one child said that he liked Dr. Suess books Sally replied, “ Dr. Suess? Me too. I’ve been reading those with (her own child). He thinks they’re pretty funny.” At the end of the assessment, Sally asked some of the students about the difficulty of the passage, and why they thought it was hard or easy. In the spring, Sally again had conversations with the students, discussing the story they had just read, giving them feedback and sharing their progress as readers as seen through the DRA results.

Sally said she learned more about her students from these literacy conversations than from the second-grade reading test. For example, referring to students’ knowledge of their own reading strategies she said that it was “interesting to see that awareness – to see if they notice what they do.” The literacy conversations that I observed in October influenced Sally’s instruction in terms of student book recommendations and the ways that she helped students learn new words.

Sally modified the second-grade test to match her instructional needs. She administered the test correctly and used it to inform her instruction. Her modifications to the test reflected a broader definition of literacy than that encompassed by the test, and her belief that assessment could be a collaborative, on-going learning process for both teachers and students.

Sarah, a third-year teacher at Webster School in Millersville district, also saw the second-grade test as an instructional opportunity. She provided feedback to the students after they had taken the test, and modified the scoring of the test to increase student motivation. After Jason finished taking the second-grade test, for example, she said to

him, “Excellent. I liked how you used your voice so I could see how there’s a difference between the two people. That made the story interesting.” To Vernon, a struggling student she said, “ You did a really nice job following along, sounding the words out.” The individualized nature of the second-grade test coupled with a structure based on the four dimensions of the test, provided a format for Sarah’s own learning as well as for feedback to the students.

In addition to providing specific feedback to students, Sarah gave strong readers “3’s”(at standard) instead of “4’s” (above standard), which was a better representation of their performance on the test. After giving Jason, a strong reader, a 3 on both phrasing and retelling she said,

I think it’s hard to give anybody a 4 for the simple fact that it’d have to be perfect and the reading that they’re at right now, they’re not gonna be, no matter how well they read.... Because if you give them a 4 now and then you go to give it to them in the spring, how are you going to show growth? So you have to give them something to work for... I think with an activity like this, giving a 4 is almost saying they can’t do any better. I think they can always do better.

Sarah modified the scoring of the second-grade test for an instructional purpose – to give her students “something to work for.” Because the QRI does not make a distinction between beginning of the year second-grade passages and end-of-the year second-grade passages, Sarah felt the need to show growth within the rubric provided by the state. She saw the second-grade test as an opportunity to learn more about her students as readers, and as an opportunity for her students to get feedback on their

reading performance. Lowering the actual score for her best readers gave them, in Sarah's mind, motivation to work harder and a way to see progress later in the year.

Teachers modified the second-grade test for many reasons. With little professional development surrounding the test, teachers were forced to rely on their background knowledge to interpret the test. In some cases, teachers' literacy beliefs were so strong that they, sometimes unknowingly, made modifications to the administration or scoring of the test. The modifications teachers made fit their own understandings of literacy and assessment and sometimes decreased the dissonance that they experienced. In other cases, when teachers modified the test to support their beliefs that the assessment could serve as an instructional episode, the teachers learned more about their students and the students learned more about themselves as readers.

#### District Factors

There were primarily two ways in which districts influenced teachers' thinking about the second-grade test: professional development and accountability.

In both Oceanview and Millersville, the professional development provided by the districts was weak and insufficient for helping teachers understand the assessment. In Oceanview, teachers were provided little guidance in terms of interpreting the new assessment, which was based on theories about literacy and assessment new to some of them. In the case of Millersville, Jean and Jody did not understand that a retelling could give them useful information that was different from the comprehension questions, and so did not administer it.

Unlike Millersville, there was district accountability associated with the second-grade test in Oceanview. As discussed in Chapter 4, district accountability led schools to allocate resources on the basis of the test. Additionally, the pressure that teachers felt to help their students perform at the grade-level standard may have exacerbated an emphasis on student scores, at the expense of viewing the test as a way to develop a more complex understanding of students. Finally, Oceanview's directives, designed to help teachers be more consistent in their administration of the test, created stress for teachers and focused some of them on one particular dimension of reading.

#### Oceanview's "Cut-Off" Numbers for Accuracy and Rate

All four teachers in Oceanview were strongly affected by a district communication about cut-off numbers for time and error. They all used a table, distributed at an early district inservice, which gave guidelines for automatically dropping students to lower levels based on a maximum time and a maximum number of errors permissible at each level. Although no longer endorsed by the district, none of the Oceanview second-grade teachers I met, including those outside this study, were aware of the district's new stance. In fact, the four teachers in my study relied heavily on the sheet to "be consistent" with other teachers in the district. The district cut-off table, which did not reflect the state's intentions behind the test and did not take into account developmental differences in the process of learning to read, created tension for teachers as they tried to negotiate a path between their own knowledge about how students learn to read and the requirements of the cut-point sheet.

For example, Sally, a teacher with 8 years of experience at Valley View, said before giving the DRA to one of her students, "I'm going to look at my 'cheat sheet.' OK, am I doing it right? I want to be sure I'm doing this test right, but with my own running records I don't worry about it. I want to have consistency." Because Sally was aware of the accountability aspect of the test in her district, she wanted to be consistent with other teachers.

However, Sally also acknowledged that one of her main concerns about the test had to do with the district's timing cut-offs. She felt that the timing cut-offs penalized children who stopped in their reading to look at the pictures or to comment on the story -- things that she valued for her students. She felt that the timing cut-offs were particularly bad at the lowest levels because, with the small amount of text at these levels, the amount of time allowed for students to read was also quite short. (At some levels this was 30 seconds.) Sally commented:

See, the problem is, if at a lower level they get stuck on a word, that's half their time. They're penalized for going back and making sure it makes sense. For going back and saying, "What is that word?" instead of just skipping along.

Sally's scoring reflected the contradictions that she felt and her concerns about struggling readers. When children recorded a time above the district cut-off because they wanted to talk about the text or were laughing, Sally decided not to penalize them and marked them at the level she thought appropriate. However, when struggling readers were even slightly over the time cut-point Sally scrupulously followed the cut sheet.

Dolores, at Oak Park, was another experienced teacher who questioned the cut-off figures. In speaking of one child who was an accurate reader with good comprehension but a slow time she said, "So this is difficult for this type of child. And I'll bet you're probably going to find this – There's a flaw in the system. In that he has very few errors and his retell was good and he can read it – there's no problem with that. That's standard, but the time wasn't standard. Does that mean ---?" Dolores initially read with Derrick at a Level 12. Even though Derrick read the story very slowly (at nearly two minutes over the cut-point), he read it accurately and gave a strong retelling. Instead of following the district cut-off guidelines, Dolores continued to read with him to a Level 18 (grade level standard) where he read slowly, accurately and with good comprehension. As an experienced teacher Dolores ignored the district directives with which she disagreed and used the test to get the information she needed.

Shirley, a third year teacher at Oak Park, was confused by the time and error cut-offs and did not understand how they had been determined. She said, "I don't understand why they allow so much time or so many errors. There's too much time. The longer passages get 9 minutes. Even the shorter passages the kids can make a lot of errors. I'll turn in the results, but for my own uses I'll use the lower levels." Although Shirley turned in her results to the district, scored as the cut-off numbers required, she nonetheless determined that she would still use her own judgment in determining how she placed students. Shirley did not understand the logic behind the cut-point sheet, and because she didn't understand it or agree with it, decided to rely on her own knowledge

to group her students. Although she lacked the years of experience of Dolores or Sally, Shirley too felt confident in using her own judgment to place students.

As a new teacher, it was difficult for Louise at Valley View, however, to ignore the district directive. Although she felt that the time cut-offs were “tough” especially for her bilingual students at the lowest levels, her first concern after listening to each of her students read was whether or not they were “in the time frame.” This strong focus tended to eclipse other aspects of the reading process and discouraged Louise from noticing other, equally important, aspects or from following what she had recently learned in her teacher education program.

Oceanview’s cut-point sheet created a dilemma for teachers between following external directives and honoring their own understandings about how children learn to read. Experienced teachers, such as Dolores, did not follow the directive when they did not agree with it, using the test to get the information about students that they wanted, even if the students were below the cut-point. Sally followed the directive when students were slow in decoding, but not when they were slow because they were enjoying the story. Shirley, a third year teacher, also followed her own beliefs in terms of assigning students to reading groups. However, my observations suggest that Louise, as a brand new teacher with a challenging assignment, simply followed the requirements and as a result was directed in how she thought about her students as readers.

The district directive caused stress for teachers as they tried to navigate a path between their own understandings of students’ reading development and the direction provided by the district. The district cut-point sheet also differentially affected the

newest teacher in this study because she lacked the time, and perhaps the confidence, to think about and modify district directives. The confusion caused by the directive also increased inconsistency in the student results that were reported to the district.

### School Factors

The school environment had a strong effect on how teachers administered, interpreted and used the second-grade test. There were four school factors that seemed most important in influencing how teachers thought about and used the second-grade test: teacher collegiality, the way the school organized for reading instruction, school-based professional development, and the principal.

### Teacher Collegiality

Teacher collegiality played an important role in how teachers thought about and used the second-grade test. Within each school, for example, the teachers administered the test in a similar manner: Louise followed Sally's lead and added interview questions to the test; Dolores and Shirley both administered the test in a short, efficient manner reminiscent of standardized tests; and Gloria and Sarah gave similar introductions. In each of these situations a newer teacher followed the lead of an experienced teacher.

Even in a school where the teachers professed not to know what the other was doing, however, the teachers administered the test in the same way. Although Jody and Jean, two experienced teachers at Lincoln School, taught across the hall from each other and had similar beliefs about literacy instruction, they had very different personal styles.

By their own admissions and confirmed by my observations, Jody talked primarily with the other adults in her classroom, while Jean interacted primarily with the other second-grade teacher. However, both of these teachers used comprehension questions instead of retellings when administering the second-grade test. They also defied the district curriculum in the same ways: continuing to teach a phased-out unit on space, pre-teaching vocabulary even when it was discouraged by the district's professional development program, and using the old basal instead of the new curriculum. They each considered themselves part of a second-grade team and were aligned with the other second-grade teacher in the building in their stance towards district initiatives.

Several teachers in the study had another teacher with whom they regularly shared thoughts about children and instruction. For example, Sally, at Valley View, job-shared with Betty (who was not part of the study). Sally taught the first part of the week, and Betty taught the last part of the week. They overlapped every Wednesday, talking about their students then, as well as frequently talking about their students on the phone during the week. The opportunity to talk about students influenced how they thought about their students as readers, and how they interpreted the second-grade test. For example, after Sally tested Wallace, Betty walked into the supply room where Sally was administering the DRA and this exchange occurred:

Sally: So I did Monica and Wallace. Wallace, I'll show you. You know, excellent. I mean he whizzed through it. It was very, very fast. He did very well.

Betty: I just read with him today. Comprehension was off.

Sally: Really? His comprehension on this was very good.

Betty: He didn't understand the nuances of the story. It's that Bony Legs, you know, the ghost comes down the chimney. So that was interesting.... And he hasn't been finishing his projects.

Sally and Betty then continued to discuss Wallace's need to move through things quickly (Sally snapped her fingers three times to illustrate this point), something that had been evident in Wallace's rapid reading on the DRA. As they talked Wallace's rapid speed began to seem a liability, not an asset. During this conversation Sally and Betty decided to hold him to completion of his projects and keep him at the same level in reading.

There were strong school effects in terms of administration of the state second-grade test at all of the schools in this study, whether or not the teachers talked regularly to each other about their practice. However, when teachers did talk to each other about their students' work on the second-grade test, as in the case of Sally and Betty, it influenced how they interpreted and used the results.

#### School Organization for Reading Instruction

How a school organized for reading instruction played a role in teachers' interpretations and uses of the second-grade reading test. Two schools, Oak Park and Webster, organized across classes for reading instruction. In both situations the instructional situation of one teacher seemed a mismatch for the second-grade test. However, the teachers at each school viewed their instructional situations very differently, partly as a result of the collegiality present between the teachers.

At Oak Park School, students were placed in one of four reading classes based on their DRA scores. As mentioned previously, two of the teachers in this study, Dolores and Shirley, expected the DRA to primarily give them information about the levels at which their students were reading so that the students could be placed in appropriate reading groups. However, one of the teachers, Shirley, saw more of value in the DRA than the identification of students' reading levels.

Shirley said that she would give the DRA even if it were not mandated because I can find out where their difficulties are. If they're able to decode... if they're having difficulty with like blends or vowels ... and of course, comprehension... If the child is fluent you can pick that up right away.

Shirley, however, taught the highest reading group (students performing above grade level) when the second grade grouped for reading. With the highest group Shirley focused primarily on comprehension and exposure to different types of reading materials "since they've got the fluency down." She did not feel that the DRA was helpful to her with these readers saying, "You know, I don't really use it." (She did, however, use it extensively during parent conferences.) Although Oceanview required several classroom-based assessments in reading ("We have so many that we have to do!"), Shirley nonetheless felt the need to create her own assessment for the top group, a writing prompt keyed to a basal text.

Although the DRA influenced Shirley's work with her homeroom students minimally (for example she prepared some extra phonics sheets for one of her homeroom students), she relied primarily on her team members to teach her lower performing

homeroom students. As the junior member of a team comprised of very veteran teachers (both of whom had over 30 years in teaching) she said, “I trust that my colleagues will do a good job making sure that they (the students) have the skills necessary to move up to the next level.”

The teachers at Oak Park shared students but there did not appear to be much conversation between them about students, or student results on the DRA, beyond the use of scores in assigning students to reading groups. Although the teachers taught each other’s students for reading, all of the classrooms seemed very separate. I observed them come into each other’s rooms after school to arrange logistics for field trips or special projects, but I never saw them discuss students with each other. During one parent conference conducted by Dolores, she told the parents of a struggling reader that they would have to come back another time to discuss the DRA and their students’ reading progress with the teacher who was responsible for his reading program. Dolores, for her part, also opposed taking time out of the student day (a regularly scheduled release time for team collaboration) to meet with the second-grade team.

Shirley, as a member of her grade-level team, used the DRA to group students. However, in general the state and district-mandated assessments were a mismatch in terms of her instructional situation. Even though Shirley valued the DRA, as a junior colleague she relied on the other, more veteran teachers, to teach reading to her homeroom students. As a result, she used the DRA only minimally to inform her instruction.

Webster School also grouped across classes for reading. However, in this situation one teacher, Gloria, taught all of the students reading while the other teacher, Sarah, taught all of the students math. These two teachers taught next door to each other, with a door open in the wall between their classes. They were constantly in each other's classrooms, talking about students. Although Gloria did not think she would learn anything from the test, after she gave it she shared her surprise with Sarah that Albert had "bombed the test" and that Julie had "whizzed right through." For her part, Sarah shared how surprised she had been at Vernon's persistence and the unusual way he had used his finger to circle each word as he read, "as if he were blocking out the other words." Gloria and Sarah also hypothesized together about why one student hadn't read in her usual baby-talk voice during the assessment. In the process of sharing their surprises they also shared possible interpretations and instructional actions.

Sarah valued the second-grade reading test saying:

I think it is a great idea. I think personally for me... the way it separated out, it really gives you concrete examples of where they (the kids) are struggling. As far as the phrasing, are they not understanding that punctuation means take a deep breath and stop? You know, is it their accuracy, are they just reading too fast and missing the words altogether? Do they need to be taught to slow down and use a bookmark? So it's (the QRI) good for skills.

Although not responsible for the main reading block, Sarah actively considered herself a reading teacher, working with her home room students in reading for a half hour each morning, reading aloud to students after lunch every day, following up on students'

Accelerated Reader daily homework, and consciously using reading and writing as a part of her math instruction. She talked frequently with Gloria about the students and their reading. Although the organization of the school day might have precluded Sarah from being active in her students' reading development, it did not because of her attitude and because her collegial relationship with Gloria helped her to keep in touch with the progress of her students during the reading block.

Although neither Dolores or Gloria said they particularly valued the second-grade test, their less-experienced colleagues, Shirley and Sarah did. Oak Park's organization for reading, coupled with low collegiality, meant that Shirley did not use the test much to inform her instruction. On the other hand, although Webster's organization for reading instruction might well have precluded Sarah from being involved in her students' reading instruction, the high level of collegiality present in that team situation meant that Sarah continued to teach reading to her students, and used the state test to help her do it.

### School-level Professional Development

With the exception of one school, school-level professional development did not play a role in the ways that teachers used and interpreted the second-grade reading test. Although the two schools I studied in Millersville both had reading cadres as a part of the Accelerated Schools program, neither of the reading cadres used the QRI results in their discussions. At Lincoln School the reading cadre focused on obtaining reading materials with a stronger sequential skills emphasis than found in the district-selected reading program; at Webster the cadre focused initially on non-fiction text and then writing. In

Oceanview District, Oak Park School opted to join the district-level literacy initiative and did not have any school-based literacy professional development. Alone among the schools in my study, Valley View developed school-level professional development which had the potential to impact teachers' interpretation and use of the state second-grade reading test.

As discussed in Chapter 4, Valley View School was in the midst of a whole-school literacy initiative during the time of my study. The school had hired a half time literacy specialist to work with teachers in their classrooms, and the staff had decided to devote all of their professional days and half of their faculty meetings to the study of literacy. Both the principal and the literacy specialist recognized the DRA as a good match with the school's literacy initiative. As Katy, the literacy specialist, noted, "It (the DRA) coincides with the work we're trying to do in the building, so it's very fortuitous." Valley's View's literacy initiative had the potential to influence teachers' understandings of the DRA in three ways. First, it created a context in which the test was valued. Because the principal and the literacy specialist valued the type of evidence that the DRA provided – teachers listening to individual students read from leveled passages – the principal, the literacy specialist and the staff considered the test good and actually better than other types of assessments.

Both of the teachers in my study from Valley View valued the test. As a result, they used it to inform their instruction, provide feedback to students, and to communicate with parents. There was one second-grade teacher outside my study at Valley View School who did not, however, value the test. The literacy specialist described this

teacher as a “resister” who thought that the test was “not good, not useful.” Katy, the literacy specialist said,

...and what I got was quite a lot of anger about this direction and this whole leveling thing and dah, dah, dah and when this school really had it together was when we all did Ginn (a basal series) and if she had her way we would go back and we would have the Ginn and la la la la la and the Ginn test would be given K-5 and we'd all be knowing what we were talking about.

This teacher's disagreement with the DRA hit at the heart of what made the DRA such a good match for Valley View and was the second way in which the school's literacy initiative had the potential to influence the teachers' use of the test. The philosophy behind the school's literacy initiative, “Know the learner. Know the strategies. Know the resources” matched a key purpose of the DRA – to help teachers understand their students individually as readers, not just as members of high, middle and low reading groups. The DRA, with its structure of many graduated text levels for beginning readers (there are 10 different levels for first grade as opposed to two levels in most first grade basal readers), had the potential to help teachers match students more closely to instructional level text. The idea of instructional level text was a key strategy in the type of individualized instruction being advocated by Katy and Jill. According to Katy, the DRA had “changed the conversation” by helping teachers to develop an awareness of “level as a concept ... more discrete levels through which children can pass. As opposed to the old ability grouping – high, low and middle.”

A third aspect of literacy instruction supported by the DRA and advocated by Valley View's literacy initiative was guided reading. (Guided reading is a method for teaching small groups of children using targeted instruction, often with instructional level text.) The principal saw guided reading as an important use for the information provided by the DRA saying, "...that's why guided reading makes so much sense to teachers because they can use that data (from the DRA) more effectively."

However, neither of the teachers in my study at that school used small group instruction to teach literacy. Sally either worked with the whole class or with students individually. She knew that she wanted to work with small groups and made the implementation of guided reading groups a learning goal for herself, but she did not know how to use the information provided by the DRA to organize for instruction. She said,

Sally: A big gap for us has been, honestly, so now what? We've got this information, now what are we going to do?

Interviewer: You mean the information from like this test and the running records?

Sally: Exactly. And now what? Do we go through each one and say, "Well, I've noticed that Joe and Susie and Kim are missing 'ed' sounds? You know what I mean? And when it's a big pattern in your class you can kind of hit it with those kids, but then it's just a few, do you just skip it and think, well, they'll get it? And that's another issue I struggle with.

Saying that she was searching for a “purpose” for guided reading groups, Sally noted that the DRA did not help her because it “just gives levels” and she already knew the levels at which the kids were reading. Sally commented, “I know the kids from reading with them. I can group them off the top of my head. But are we missing something?”

Sally and her teaching partner, Betty (who was not in the study) did get some help in thinking about guided reading groups, although not as a part of their classroom instruction. As a result of a grant, Sally and Betty tutored four of their struggling readers after school. Katy helped them build on the model of Early Literacy Groups (small group tutoring built on a Reading Recovery model, Clay, 1993). This organization for reading instruction shared the same theoretical base as the DRA and had guided reading as a key component. Sally noted that they planned their instruction by “...looking at what they (the students) were missing as far as what good readers do. And when it was only four (students) it was easy to see this.” Sally and her teaching partner had also planned to implement guided reading in their classroom using the new federally-funded teacher to oversee their class while they worked with small groups. However, this plan fell through due to philosophical and political differences with the new teacher. (She felt her job description required her to work only with struggling readers.)

Although she was a knowledgeable and experienced teacher, able to use the DRA to inform her instruction of individual students, Sally was unable to use the reading dimensions provided by the state second grade test or the more refined instructional levels of the DRA to help her plan and carry out small group work in her classroom. She

needed help, in terms of professional development and in-class support, to learn how to use the information provided by the test to determine appropriate instructional targets and implementation.

Because of the professional development available at her school, Sally received the help she needed. However, philosophical and political differences between staff members prevented her from putting guided reading into place in her class. Louise, a brand new teacher, lacked both the resources and the expertise to put a more complex individualized reading scheme into place. Although the professional development offered by their school was helpful to both of them, their cases illustrate the depth of professional development, support, resources and time necessary in order for teachers to implement fundamental change in practice.

### The Role of the Principal

School principals primarily played a role in how teachers implemented and used the test. The principals influenced how teachers felt in several ways: by providing release time or other support for the test, by mediating district pressure, by conferencing with teachers, and by providing students and teachers with resources so that they could be successful.

At all of the schools, the principals either helped the teachers in some way to administer the test, either through release time or supplemental staffing, or teachers' classrooms were well-enough staffed that they did not need additional resources. When principals did provide extra support, this made a positive difference in teachers' attitudes

toward the test. For example, one of the newer teachers, Sarah, who was at Webster School, noted that she liked the state mandate because as a result her principal gave her release time, in her mind a necessary requirement for giving the test.

Support for giving the test seemed especially important for new teachers. Shirley, a new teacher, and Dolores, a veteran teacher, both gave the test during the fall without release time. (They used release time in the spring.) Administering the test seemed difficult for Shirley, a new teacher, to manage. She tried to test all her students during the short periods when her class was with a specialist teacher (such as PE or computers). Testing did not appear difficult at all for Dolores, a veteran teacher, who administered the test to individual students while the rest of the class worked quietly at their desks. Although Dolores complained about administering the test in the fall without release time, she also said, “But teachers often test children during class time – it’s not an unpopular thing – it’s been done for years.”

Although all of the teachers in the study were offered a half-day of release time to test their students and talk to me, none of the teachers took advantage of this offer, saying it was more trouble than it was worth. Additionally, almost all of the teachers extended the test beyond the state or district requirements. By supporting teachers in doing their own testing, principals sent a message that there was value in giving the assessment. This action also acknowledged the extra work that teachers were undertaking.

The ways in which principals mediated between the district and the teachers also made a difference to teachers’ attitudes. For example, Jill, the school principal at Valley View, buffered the pressure that teachers were feeling to get their students to achieve at

the district standard. Sally noted that the increasing pressure being applied to teachers by Oceanview district was “scary” and “really undermining.” She commented that she and her teaching partner “talked about it (the increasing pressure) all the time.” Nonetheless, they were firm in teaching students in the way that they thought was best saying, “we don’t really let the district tell us, and we don’t really let the school say.” Sally added that it was her principal who made this stance possible, stating, “I just choose to ignore it (the district pressure) and go on my way. As long as I can. And Jill (the principal) lets me.”

At Valley View the teachers were personally held accountable for getting their students to standard. Sally, however, seemed to relish this challenge, noting that in the three years she’d been predicting which students would get to standard, she’d always made her goal. Louise, a first year teacher, felt anxious about being held to this standard, but at the end of the year nearly all of her students had met the standard, thanks to her own determination as well as the resources that had been shifted her way by the principal. The principal’s support of these teachers, both psychologically and tangibly, enabled an experienced teacher to see her personal accountability as a challenge, while enabling a new teacher to successfully meet the challenge.

At Valley View, Sally found it helpful to meet with the other second grade teachers and the principal to discuss their students’ DRA results and to plan for the use of resources. At Lincoln School in Millersville district, Jody and Jean also welcomed individual monthly conferences with their principal Mike. It was during one of these

conferences that Jody and Jean each shared their students' QRI results. In regard to sharing the results, Jody said about her principal:

Oh, he's incredible. He meets with us like every month... and we go down on every student in the class and how they are academically, emotionally, socially, behaviorally – anything that they need. So he's clear on where they are in every academic area. Everything about them. And that's when he looks at it (the QRI results)... He takes copious notes and you say, oh this kid needs whatever, he's got the social worker in or whatever. And he is down on the ground talking to him (the child) saying, "What are you doing? Why are you doing this?" He's just so involved.

Shirley, at Oak Park in Oceanview district, commented that her principal had used the second-grade test to do a "survey on ethnicity and all that and brought the scores back to us ... which was alarming because it was a big percentage of our African Americans who were at a lower level. But now they have moved up!" She saw this use of the test, as a tool to reflect back to teachers the patterns of achievement in their students, as worthwhile. She also felt that the test gave her principal legitimate reason to question teachers on their practice saying,

Our principal has actually focused on making sure that the kids that are not quite meeting the level are provided with a lot more extra help this year... Now she's got something she can go by and can come up to us and say, "Hey, this child's not meeting the standard so what are you going to do?"

However, Dolores, a veteran teacher with many years of experience in the district, thought that the district was unclear about its purposes for the second-grade test, did not have a congenial relationship with her principal, and did not value the test, was suspicious when her principal asked for her DRA results, saying, "I don't think it's clear to teachers why they're using this test... Whether it's going to be used as a teaching tool or is it going to be used to evaluate principals – because our principal asked for our test scores and we're not sure what that's all about." Dolores' own suspicions of the test and her principal, and perhaps her earlier experiences in the district, led her to question the actions that Shirley admired.

There were three schools in this study in which the principals used the results of the second-grade test. In all of these schools, the result of most of the principals' actions, whether in terms of physical resources, psychological support or reflecting teachers practice back to them, was to help school staff work together to support student learning, which encouraged the development of a collegial atmosphere.

The school environment was an important factor in the ways that teachers thought about and used the test. There were strong school effects in each school in terms of the way that teachers administered and used the second-grade test. Furthermore, the degree of collegiality at each school seemed to positively influence teachers' attitudes, interpretations and uses of the test, especially in team-teaching situations where teachers shared student results or when the test was linked to school-based professional development. In some cases collegiality was able to ameliorate situations that might have promoted minimal use, such as organizing reading groups across classes, or negative

attitudes, such as holding teachers directly accountable for their students results. School collegiality, the way that schools grouped for instruction, professional development, and the principal all had the potential for influencing teachers' use of the test.

### The Value of the Second-Grade Test

All of the teachers in this study thought that the primary reason the state had mandated a reading test was to create accountability at the second grade level as a prelude to the WASL. As Sarah, a teacher at Webster School said,

I think you have to have some sort of paperwork to show that teachers are teaching reading, how they're assessing it, and what they're doing to improve it, if it's below grade level, especially (because of) the WASL in grades 4, 7 and 10.

You know there has to be some sort of accountability ... This child was shown in second grade that he was below average, what has each teacher done subsequently to help him out?

The teachers were not opposed to this accountability, however, and in most cases seemed to welcome it because the state second grade test provided authoritative evidence that their students had made progress as readers. Teachers devoted more attention to the dimensions of the test, and used the results with students, with parents and with outside audiences such as the school board. Most of teachers in this study supported the state second-grade reading test in some way.

### Authoritative Documentation of Student Growth

The ability of the state second-grade test to document student growth was the most important reason that teachers valued this assessment. As Shirley, a teacher at Oak Park said, when she explained why she was going to administer the DRA to all of her students before the district required this,

Just to let them (the students) know where they are. And to let me know... Not just OK, they're at level; I want to see how much progress they've made. And if they haven't, I need to know that and the parents need to know that. I mean, that's what it's all about, right?

The desire to document student growth in reading, either for the school board, parents or themselves, was the major reason that most of the teachers (with the exception of Dolores at Oak Park) administered the test beyond the requirements of the state or their district. For example, Oceanview teachers planned to administer the test to all of their second graders in the spring before it was required by the district, Millersville teachers assessed all of their students to their instructional level at least twice a year, and Sarah and Gloria, who taught multi-age Grade 2/3 classes at Webster, assessed their third graders in the fall. (They planned to test third graders in the spring as well but ran into logistical problems.) This expansion of the test by almost all of the teachers was particularly striking because most of the teachers also complained about the time-consuming nature of the test.

The outside validation that the test provided was also important to teachers. Jody, who taught at Lincoln School and who claimed that she did not get any new information about her students from the QRI, nonetheless assessed all of her students using the QRI

three times during the year. She and Jean, another teacher at her school, had a strong reason for documenting student growth – they wanted to convince the school board that the pilot program (which put a full-time aide in each second grade classroom) should be re-funded. Jody chose to use the QRI as one of the key pieces of evidence for the board presentation because as a “state assessment” it was a “stronger piece” of evidence than their other assessments. (As Jody said, referring to the basal tests she commonly used, “It’s not just an HBJ end-of-unit test.”)

The outside validation that the test provided was also important in terms of reporting to parents. Sally, who taught at Valley View, said,

I think Betty (Sally’s teaching partner) and I do know our book levels ... I don’t know why I have to believe in someone else – maybe because someone is telling me this is Level 24 and I’m saying, “OK.”.... Well, and we share our information with parents and it just kind of gives us a leg to stand on and say, “Wow, your kid reads at second grade. This is great!”... It just makes it more official.

All of the teachers in the study felt some pressure to have their students do well on the second grade test. Louise, the newest teacher in my study, who was working at Valley View, a low-scoring school in Oceanview, felt the pressure most acutely, however. Although she was anxious about helping all of her students meet the standard, by the year’s end all but one of Louise’s students had met the second grade standard. The second grade test validated the progress that her students had made and, after administering the DRA for the second time, Louise said, “I didn’t realize that they had made such great gains!” Louise took pride in her students’ achievements and used the

DRA to show them the progress they had made. She described her conference with Jason by saying

“Jason,” I said. Let’s look. You read this in the fall and now you can read this!” and I opened the books (the DRA leveled books) and I said, “Lookit! Isn’t that amazing!” Half the kids I was crying with. You know. I wanted them to visually see... It was cool!

### Teaching to the Test

All of the teachers, with the exception of Louise, said that they taught students based first on what they thought students needed, not on the basis of outside standards or tests. However, it appeared that all of the teachers were influenced by the second-grade assessment in some way. Shirley, who taught at Oak Park, stated this clearly when she said, “...now we know what we are working towards which I think is helpful to me to make sure that these kids are at that level. And again, this puts a little more pressure on us, too.”

Even teachers who disagreed with the test, or didn’t learn anything from it, modified their instruction so that their students could meet the standard. Dolores, who taught at Oak Park and was highly critical of the test, nonetheless focused on rate to such an extent with Derrick, a child who read accurately but slowly, that he said to her during the spring administration of the test, “Am I reading fast enough? Am I doing OK?” Jody, who taught at Lincoln and who did not get useful information from the QRI, administered it three times to her students, using it as a benchmark against which to judge her students’

achievement. Although Jody said the core of her instruction (phonics, comprehension, “process, process, process print,” and ownership) had always been there, it was hard not to wonder if the emphasis on rate and phrasing that I saw during my observations (for example students re-reading basal stories to “work on fluency and rate”) was influenced by the test.

All of the teachers in this study actively taught to the dimensions of the second-grade test. As a result, in all classrooms in this study, all students with the exception of those who were new speakers of English or had a special education designation met the second-grade standard.

#### The Second-Grade Test as a State Mandate

At the end of this study I asked teachers if they would give the DRA or the QRI if it were not mandated. Three of the teachers – Dolores (in Oceanview), Jody and Gloria (in Millersville) – said “No.” All of these teachers felt that their own assessments were preferable to the state mandate. In addition, Dolores felt that the state testing was a “knee jerk reaction” and the testing was a substitute for the real issues of class size and teacher pay. She also thought that the various state and district mandates disregarded the professionalism of teachers like her saying, “People with my background are not honored for what we know and can do.... Teachers are entrepreneurs. I know what’s best for me.” Jody had a slightly different opinion when she said,

And I think it’s fine to have accountability. We really do need something to base it on, but then what are we going to do with it if they don’t meet the standard...

For the kids that don't cut it, they're not giving any solutions as to how to get them up to snuff but "Get them there!"

For her part, Gloria felt that the test was good "state-wide" so that resources could be distributed to students who needed them, but that it was not necessary at her school where assessments and resources were already in place.

Four of the teachers, Sally, Louise, Shirley (in Oceanview) and Jean (in Millersville), said that they would continue to give either the DRA or the QRI even if it was not mandated, because of the way that it documented student growth and gave them information about individual students. To my surprise, however, three of the teachers -- Louise (Oceanview), Jean and Sarah (Millersville) -- liked the fact that the second grade test was a state mandate. For Sarah, the state mandate meant that her principal provided release time to give the test, which she thought was a necessary component for administering it. For Jean and Louise, the mandate meant that what they particularly valued about the test -- the opportunity to listen to students read one-to-one -- was sanctioned by the state and as a result would not "slip," as Louise said. Jean said, "I think it's very worthwhile. It's good for teachers to have to do because I think it's easy not to do it. Not to take that time to really individually assess. Sometimes it really has to be mandated."

Six of the teachers in the study supported the second grade test to some degree, which is unusual for a mandated state testing policy. The teachers supported the policy because, as a state mandate, it authoritatively provided documentation of student growth in reading and validated their success as teachers. At the end of the year in the classes in

which I observed, all students, except those with special education designations or who were just learning to speak English, met the second grade standard. As Sally, who taught at Valley View said, the test was “affirming” because “it’s really encouraging to see the growth! Wow! We’re doing something right!”

### Summary

Teachers in this study learned from the test if they expected it to deepen their understandings of children as readers or if they expected it to help them place students on a continuum of reading achievement. Teachers also learned when there was dissonance between their expectations for student performance and the ways that students actually performed on the test. Teachers’ beliefs about literacy and assessment influenced how teachers thought about the test, and also caused teachers to modify parts of the test. Some of the modifications decreased the dissonance that teachers experienced and limited what teachers learned from the test. Other modifications to the test extended teachers’ learning.

District and school contexts also played a role in how teachers interpreted, used and felt about the test. In particular, the lack of professional development offered by districts forced teachers to rely on their previous understandings of literacy and assessment, while an emphasis on accountability encouraged teachers to primarily view the test as a way to measure student performance as opposed to helping teachers develop a more complex understanding of reading development. There were strong school effects in this study, with similarities in administration and use found among teachers in

the same school. Teachers were influenced in their thinking at the school level by the level of collegiality, the way the school organized for instruction, school-based professional development, and the way that the principal used the assessment.

All of the teachers in this study, whether or not they were in districts that made teachers accountable for their students' results on the test, felt pressure to help their students meet the state standard. As a result they all taught to the dimensions of the test, and at the end of the year had high success rates among their students. More than half of the teachers liked the test because it provided authoritative evidence of student growth in reading, and authoritative evidence that they had helped their students meet the state standard.

## **Chapter 6**

### **Discussion**

The purpose of this chapter is to discuss the findings and implications of this study for practice and further research. This chapter also explores the limitations of the study.

This study examined how eight second-grade teachers understood, used and valued a state-mandated reading assessment, and the ways in which the teachers' understandings were mediated by school and district contexts. The chapter begins with a discussion of the findings, using the themes of the questions that guided this study as organizers. These sections are followed by the implications of the study, the limitations of the study and a final conclusion.

### Findings

#### Understanding and Using the Second-Grade Assessment

Teachers in this study understood the second-grade test through the lens of their own literacy and assessment beliefs and practices. The results are congruent with earlier studies that show that curriculum reform policies are enacted in the context of teachers' existing beliefs and practices (Cohen & Ball, 1990; Spillane & Jennings, 1997).

The beliefs that teachers held were strong. Teachers modified the administration and scoring of the test to fit their beliefs, even in the face of dissonant results or previous professional development. When teachers made modifications to the test based on assumptions about literacy or assessment not shared by the test, the dissonance that they

experienced between the test results and their previously held expectations of students was decreased. As a result, these teachers learned less about their students and their practice. In contrast, some teachers modified the test in ways that revealed a more sophisticated understanding of reading than was represented by the test instrument itself. In these cases the teachers learned more about their students and, in some cases, the students learned more about themselves as readers.

This finding has several implications for policy. On the one hand, to the extent that teachers perceived the policy fitting into their existing practice, the teachers could support and embrace it. A policy can inform teachers about their practice, but only if they allow it to do so (Ball & Cohen, 1990). On the other hand, if the policy does not hold out an elaborated or different vision to teachers, there is nothing to learn. For some of the teachers, the use of four dimensions to evaluate students' reading ability, instead of a single composite score, provided a more expanded vision of reading assessment and instruction than they had previously experienced, and thus they learned something new about their students. However, where the gap was too great between teachers' literacy or assessment practices and the test, teachers' beliefs hindered their ability to learn from the policy. This finding speaks to the need for contextualized professional development – professional development that mediates teachers' understandings by offering them the opportunity to reflect on their students' results in the light of their own practice.

There were several aspects of the assessments themselves that helped teachers learn more about their students as readers. In the first place, because the second-grade assessment was classroom-based, teacher decision-making was at the heart of the

reporting process. As a result, teachers were provided with an opportunity to reflect on their students' reading development and the nature of the reading process. Second, the structure of scoring encouraged teachers to focus on various dimensions of reading instead of a single grade-level score. Third, the provision of materials for assessing readers at a variety of levels encouraged teachers to assess students at a range of levels. And finally, teachers saw the assessment as a valid representation of the reading abilities of young children and as a result were more willing to seriously consider the results.

#### Valuing the Second-Grade Test

All of the teachers in this study believed that the primary purpose of the state reading assessment was to provide accountability at the second-grade level in preparation for the fourth-grade high-stakes WASL. Most of teachers welcomed the accountability that this test provided because it provided authoritative evidence that their students had met the second-grade reading standard. The teachers used this evidence to convince themselves, as well as parents, administrators and school boards, that they had prepared their students adequately for success on the high-stakes fourth-grade test. Because they valued the "official" documentation of student progress that the test provided, most of the teachers administered (or planned to administer) the test beyond state or district requirements.

That most of the teachers administered the test beyond the state or district requirements, despite their complaints about its time-consuming nature, was stunning to me. Teaching is an uncertain craft (Lortie, 1975; McDonald, 1992). In an era of

increasingly high-stakes, when the uncertainty of teaching is often ignored by policy-makers and district-office administrators, teachers crave proof that what they are doing is sufficient, so that they can say, as Sarah did, “I actually did teach your child something.” The high-stakes environment intensifies the need for teachers to know that they have made a difference for children.

Even in districts that did not emphasize accountability to the second-grade assessment, all of the teachers wanted their students to meet the second-grade standard by the end of the year and so taught to the dimensions of the test. This was helpful, in that teachers were aware of the dimensions of reading, but also problematic in that it may have narrowed teachers’ instruction. The second-grade assessment may have become *the* benchmark of student progress for some teachers instead of one of many indicators.

Most of the teachers used the assessment as a “locator device” (Shepard, 2000a), either to assign their students to ability-based reading groups or to validate that their students were reading in appropriate level text. The two uses of accountability and locating focused teachers’ attention on the test as a measure of performance, which hindered teachers in also using the test as a tool for developing knowledge about students as readers. The perceived accountability demands of the test, coupled with teachers’ previous understandings of assessment, sometimes hampered the policy in fulfilling its educative function.

The second-grade test seemed especially important to new teachers. Two of the new teachers spontaneously said that they liked the state mandate, while the third new teacher said she would continue to give the DRA even if the district did not mandate it.

For these teachers, the state test provided an opportunity to connect with students individually and an opportunity to share documentation of progress with students and parents. The new teachers' recent completion of a teacher education program may also be a factor in the way that the teachers viewed the test, as the newer teachers all had a belief in the importance of individual assessment. These teachers had a different view than some of the experienced teachers whose assessments primarily influenced group as opposed to individual instruction, and who viewed assessment mainly as summative, as illustrated by their reliance on basal reader end-of-chapter tests and Accelerated Reader tests. Some experienced teachers perceived the test as less important because they had previous assessments in place that they felt supported their instruction. However, for the most part, the assessments that these teachers had in place gave different information than that provided by the second-grade assessments.

For the newer teachers, and for those experienced teachers who had an orientation towards individualized instruction, the test also provided important symbolic validation of a practice they held dear – assessing children individually (Malen & Knapp, 1994). This validation of this practice, coupled with the structure of the test that emphasized reading complexity, provided an opportunity for teachers to develop important knowledge not only about individual students, but also about reading development. This in turn, had the potential to help teachers develop pedagogical content knowledge (Carpenter, Fennema, Peterson & Carey, 1988). Understanding how individual children think can also provide a foundation for teachers to make fundamental changes in instruction (Fennema, et al., 1996)

### The Mediating Role of Districts and Schools

Districts and schools played an important role in how teachers in this study understood, used and valued the state assessment. In line with previous research (Hertert, 1996; Spillane, 1996; Standerford, 1997), this study found that a district's ability to support a state policy hinged on the compatibility of existing district policies as well as the district's capacity to muster the necessary resources. In this case, district support also hinged on whether or not the state assessment met a felt district need, as it did in the case of Oceanview, or whether it was seen as an addition to an already existing district assessment arsenal, as it was in Millersville. Oceanview used the state policy to legitimize its changes in district assessment policies. As a result, it proliferated new assessments and district policies in response to the state policy, as Spillane (1996) predicted would happen in a high capacity district whose policies were aligned with those of the state. On the other hand, Millersville, a small capacity district, did what was minimally necessary to meet the state requirement. As a result resources were conserved and there were minimal disruptions at the district level, as Standerford (1997) would have predicted. However, while it was "business as usual" at the district level, it was not "business as usual" at the classroom level, as teachers administered the test and in most cases learned from it.

Since this policy came with little state support, district involvement was particularly important to its success. There were two aspects of district support that seemed important to the success of the policy. One was the degree of accountability that

districts attached to the assessment. This had positive and negative outcomes. On one hand, the high degree of accountability that Oceanview attached to the second-grade assessment encouraged allocation of resources to students who needed help. However, the degree of accountability also focused district attention on attempting to insure standardization in test administration and scoring. In Oceanview, district directives to encourage teacher consistency caused stress for teachers because the directives conflicted with their understandings of reading development. Oceanview's directives also focused on teachers as deliverers of test results and ignored the central role that teachers' knowledge and judgment played in the administration and scoring of the test. The district emphasis on accountability may also have further prompted teachers to focus on the achievement levels of their students and less on understanding student performance and implications for instruction. Thus, Oceanview's emphasis on accountability, while helping students get the resources they needed and focusing teachers' attention on the test, at the same time had some negative consequences for teachers. It helped to create a climate in which it was more difficult for teachers to learn about their students as readers. This case illustrated the complex interactions that occur when tests are used for the dual purposes of accountability and for informing teachers' instruction.

On the other hand, Millersville, with little district accountability attached to the test, placed little emphasis on the assessment. As a result, resources were not distributed on the basis of the test results. However, two of the Millersville second-grade teachers created their own high-stakes accountability when they used the scores from the second-grade test to report to the school board. At the other Millersville school, which had a

strong school-wide emphasis on the WASL but no emphasis on the second-grade test, second-grade teachers still felt a responsibility to help their students meet the standard. District pressure did not play a role in the accountability that these teachers felt. They chose to use it as a part of their own self-imposed accountability because the test seemed valid to them or because it carried the authority of a state mandate.

The cases of Millersville and Oceanview raise the question of the utility of district pressure on teachers. In Oceanview, district pressure to have children meet the state standard on the second-grade assessment had negative and positive consequences. On the one hand, district pressure gave the test importance and played a role in its broad use at the school and individual teacher level. On the other hand, district accountability directives, particularly the number of changes in directives over time and the focus on scores, led teachers to be concerned with standardization and as a result created stress and hindered teacher understanding. In Millersville, district pressure was not necessary to create feelings of accountability on the part of teachers. However, the teachers at one school in Millersville did fundamentally alter the test in a way that was consistent with their more limited view of reading. The lack of district pressure made it easier for teachers to change the test, and as a consequence they learned less from it. This study suggests that district pressure promoted concern with standardization, both good and bad. However, regardless of district pressure, teachers wanted their students to do well on the test.

Professional development was the second key aspect of district support that influenced how teachers interpreted and used the test. Because the focus in both districts

was on procedural professional development as opposed to substantive professional development (Spillane & Thompson, 1997), teachers learned how to administer the test without much understanding of what it meant or how it could be used to help them teach. As a result teachers interpreted it according to their previous understandings about literacy and assessment, which were sometimes in conflict with the intent of the state test. This was especially damaging in Oceanview, where the assessment chosen for the second-grade test was based on different assumptions about literacy learning and assessment than that held by some teachers.

School contexts also played a role in the ways that teachers understood the test. The results of this study suggest that there were strong school effects in how teachers administered and used the tests, even in schools with seemingly low collegiality. These findings were similar to those of Hoffman and his associates (Hoffman et al, 1998) who found strong school effects in the ways in which teachers used basal readers. The collegial nature of the school, which was manifest either informally or through the provision of school-level professional development, also affected the opportunities that teachers had to learn from the test, as would be expected from previous school-based research (Hamilton & Richardson, 1995; Little, 1987; Rosenholtz, Bassler & Hoover-Dempsey, 1986). However, much of the school-based research on teacher learning focuses on teachers' participation in voluntary professional development activities, while this study was focused on teachers' responses to a mandated state policy.

The principal, in providing resources and personal support, was also pivotal in helping teachers see the second-grade reading assessment as an opportunity for all school

personnel to work together to help students be successful readers. However, as Barbara Scott Nelson argues, principals also “enact, daily, a set of ideas about the nature of learning and teaching, thereby influencing the intellectual climate of schools in particular ways” (Nelson, 1999). The second-grade reading assessment represented a vision of assessment and literacy different than that held by some teachers. This vision viewed assessment as individual and literacy as complex. When the principal understood these dimensions, his or her own practice of the principalship was affected. For example, when a principal understood the link between individual assessment and targeted instruction, school-level professional development and teacher observations were more likely to focus on this issue, increasing the potential that teachers would develop this understanding as well. Conversely, when a principal primarily valued the numerical score of the test, the complexity of the reading process was hidden and most likely not discussed as a part of school-level data analysis. How principals understood the dimensions of the second-grade assessment had the potential to influence how teachers understood and used the assessment.

### Implications for Practice and Research

There are several implications from this study for practice. First, the state second-grade reading test provided a “teachable moment.” The test impacted how teachers thought about individual students and the reading process. However, without accompanying professional development, the potential of the test to inform teachers about their students, their instruction and the reading process was not fully realized.

Given the important role that collegiality can play in developing teacher understanding, and the situated nature of that understanding, professional development should provide teachers the opportunity to discuss the assessment with colleagues in light of their own practices and their own students' results.

The ways that teachers interpreted the second-grade reading test was based on their previous understandings about assessment, as well as literacy. Therefore, professional development about assessment practices, and in particular about the underlying concepts reflected by the test, is important. Furthermore, the different understandings that teachers had of the test, especially with regard to experienced and new teachers, argues for several types of professional development offerings. For example, it may be important to offer professional development to veteran teachers who are not facile with current ideas about reading, which recognizes their knowledge and experience, but which at the same time creates dissonance in their thinking. Acknowledgement in combination with dissonance has the potential to create a situation that helps teachers consider new ways of thinking about instruction and assessment. Similarly experienced teachers with a strong understanding of individual assessment and literacy processes would likely value professional development that helps them reflect on students and pedagogical content knowledge in the context of their own practices. In the situation of new teachers who have an understanding of the importance of individual assessment, professional development which helps them use and more deeply understand literacy processes and evidence of student development might be most helpful. Most

important, this study reminds us that fundamental change in practice takes substantial time and resources (Shepard et al, 1996).

Second, the school context, particularly the principal and the nature of collegiality around teaching and learning, made a difference in how teachers perceived the test. As a result, literacy professional development should include the principal. The results of this study also emphasize the potential of school collegiality to deepen teachers' understandings of literacy and assessment, especially when teachers share students. However, the strong school effects shown by this study also argue for the importance outside expertise to expand teachers' knowledge.

Third, even without strong accountability measures in place, teachers felt pressure to help their students meet the second-grade standard as measured by this test. Almost all of the teachers valued the documentation that the test provided that their students were on track towards meeting the fourth grade standard. It was apparent that although teachers felt pressure from the high-stakes environment of the fourth-grade state test, they also felt accountable to students, parents and themselves. In this environment, low-stakes classroom-based assessments can help teachers modify their instruction and improve student learning, especially when teachers fundamentally agree with or come to understand the assumptions underlying the assessment.

There are also implications from this study for further research. There is little research on teachers' understandings of high-stakes performance assessments; there is also little research on how teachers use and understand classroom-based assessments. Collaboration between classroom teachers and researchers would be an especially useful

form of research in this area, as it would generate improvement in classroom practice at the same time it advanced knowledge in the field (Shepard, 2000b). Furthermore, since this study confirms the importance of district and school contexts in influencing teachers' implementation of state policy, more research on the impact of state policy at the classroom level in light of these contexts is necessary if we are to understand effects of state policies for teachers and students.

#### Limitations to the Study

There were several limitations to this study. First, there were only eight teachers in the study, representing four schools and two districts. This is a small number of teachers, working in particular district and school contexts, and so the results may not generalize to other teachers. Furthermore, most of these teachers had support to administer the test, and so they might well have a more positive attitude towards the assessment than teachers who were not supported with release time or supplemental staff.

Second, all of the teachers in this study were volunteers. It was difficult to find teachers who were willing to participate in this study – I almost gave up on following the original study design – and so not only are the teachers volunteers, but they are also more likely to be teachers who take risks and who are most confident in themselves and their practice. They may also be teachers who felt most comfortable with the assessment, giving the study a more positive spin than might otherwise be expected.

Third, I gained access to schools in Oceanview because of my friendships with the head teacher in one school and the literacy specialist in the other. Because of my

previous relationships with these people, the data I collected in Oceanview was rich. These participants were not hesitant or reserved in their answers to my questions. Due to their positions midway between the principals and the teachers, they were also knowledgeable about conflicts in the school and district, and because they had both been at their schools for several years, they were cognizant of school and district history in relation to literacy and assessment. The participants volunteered details about the inner workings of the schools and district that were not obvious on the surface, and they were kinds of details that I did not begin to find out until the end of my observations in Millersville. The friendships and trust that I had with these people also eased my initial encounters with most of the teachers in the Oceanview schools, and I think shortened the time it took to develop rapport and trust with them. Therefore, I have richer data from the schools in Oceanview than the schools in Millersville, where I did not personally know anyone in the district before beginning the project.

Fourth, due to my difficulty in finding participants and the short window of time in each district when the second-grade tests were administered, the think-alouds were my first observations of the teachers. It would have been far preferable to have interviewed the teachers and observed one time in each of their classrooms before doing the think-alouds. Although I tried to be as neutral as possible, I know that my observation of the newer teachers administering the test was stressful for them. Based on my later conversations with teachers, it is safe to assume that they were more reserved in their comments during the think-aloud because they didn't know me. Also, if I had interviewed teachers and observed in their classrooms prior to the think-aloud I would

have had some understanding of their instructional stances and as a result would have done a better job of asking probing questions.

Fifth, I was also the only observer in this study. While I tried to create several data points and some checks on my own biases (member checks), this study clearly reflects my own perspective. Although the observations and interviews have consistency as a result of being collected by one person, over time my perspective may have changed, with one set of teacher observations influencing others. I tried to counter this by spending a minimum of four hours in each classroom on at least two separate occasions, and by collecting observations and interviews over the course of the year, which deepened my understanding of each individual teacher and classroom. Semi-structured interview protocols and my background as a former second-grade teacher and reading specialist who had administered many DRA and QRI assessments to students, also made me less susceptible to confounding the data between classrooms. My background in this area helped me to understand the data I was collecting; on the other hand I had to guard against the trap of confusing familiarity with understanding. I also had to be careful not to infuse my personal perspectives on good practice into my interviews, observations and analysis.

Finally, there is a strong possibility that my presence influenced the results of the study. Although I purposely designed the study to occur within the context of natural school situations, nonetheless this study gave teachers the opportunity to reflect on the assessment that they most likely would not have otherwise had. I asked one teacher if she thought that the study had influenced her thinking, and she replied that it had perhaps

encouraged her to analyze her results more than she would have without my interest. It is also possible that in Millersville, where the assessment pressure was minimal, the study added importance to the test.

### Conclusion

The results of this study suggest that the Washington State second-grade reading assessment was successful in its goal of helping struggling readers obtain the assistance that they needed, while also providing useful information to teachers. Although different teachers implemented the assessment in different ways and received different levels of professional development, it did help most teachers modify their instruction to support the reading development of their students. The test was also used by all teachers to report to parents, as required by the legislation. Two schools in the study also used the test as the basis for the distribution of resources to struggling readers. As a low-stakes test, it also appeared to have few of the negative associations of high-stakes tests. Many factors contributed to the success of the policy including the fit between teachers' beliefs and the assessment, supportive school contexts, and the nature of the assessment instruments themselves. This study suggests that it is possible to craft a state assessment policy that teachers can learn from and support.

## References

- Airasian, P.W. (1988). Measurement driven instruction: A closer look. Educational Measurement: Issues and Practices, 7, 6-11.
- Aschbacher, P. (1994). Helping educators to develop and use alternative assessments: Barriers and facilitators. Educational Policy, 8, 202 - 223.
- Ball, D. (1990). Reflections and deflections of policy: The case of Carol Turner. Educational Evaluation and Policy Analysis, 12, 241-245.
- Bereiter, C. & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L.B. Resnick (Ed.), Knowing, Learning and Instruction: Essays in Honor of Robert Glaser (pp. 361-392). Hillsdale, NJ: Lawrence Erlbaum.
- Botsford, T. & Ryan, S. (1998). Second Grade Reading Testing Program Participant Handbook: Implementation of ESHB 2042. Olympia, WA: Office of Superintendent of Public Instruction.
- Carpenter, T.P, Fennema, E., Peterson, P.L., Carey, D.A. (1988). Teachers' pedagogical knowledge of students' problem solving in elementary arithmetic. Journal for Research in Mathematics Education, 19, 385-401.
- Cohen, D. (1990). A revolution in one classroom: The case of Mrs. Oublier. Educational Evaluation and Policy Analysis, 12, 311-344.
- CTB/McGraw Hill (1990). Comprehensive Tests of Basic Skills. Monterey, CA: CTB Macmillan/McGraw Hill
- Cuban, L. (1993). How teachers taught: Constancy and change in American classrooms, 1890-1980 (2<sup>nd</sup> ed.). New York: Teachers' College Press.

Erickson, F.(1986). Qualitative methods in research on teaching. In M.C. Wittrock (Ed.), Handbook of research on teaching 3rd edition (pp. 119-161). New York: Macmillan.

Falk, B. & Ort, S. (1998). Sitting down to score: Teacher learning through assessment. Phi Delta Kappan,80, 59-64.

Fennema, E., Carpenter, T.P., Franke, M.L., Levi, L., Jacobs, V.R., Empson, S.B. (1996). A longitudinal study of learning to use children's thinking in mathematical instruction. Journal for Research in Mathematics Education, 27, 403-434.

Firestone, W.A., Mayrowetz, D., Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. Educational Evaluation and Policy Analysis, 20, 95-113.

Glesne, C. & Peshkin, A. (1992). Becoming qualitative researchers: An introduction. White Plains, NY: Longman.

Hamilton, M. & Richardson, V. (1995). Effects of the culture in two schools on the process and outcomes of staff development. The Elementary School Journal, 95, 367 - 385.

Heaton, R.M. (1992). Who is minding the mathematics content? A case study of the fifth-grade teacher. Elementary School Journal, 93, 153-162.

Hieronymus, A.N., Hoover, H.D., Oberley, K.R., Cantor, N.K., Frisbie, D.A., Dunbar, S.B., Lewis, J.C., Lindquist, E.F. (1990). Iowa Tests of Basic Skills. Chicago: Riverside Publishing Co.

Hertert, L. (1996). Systematic school reform in the 1990s: A local perspective. Educational Policy, 10, 379-398.

Hoffman, J.V., McCarthey, S.J, Elliot, B., Bayles, D.L., Price, D.P., Feree, A., Abbott, J.A. (1998). The literature-based basals in first-grade classrooms: Savior, Satan, or same-old, same-old? Reading Research Quarterly, 33, 168-197.

Hollingsworth, S. (1989). Prior beliefs and cognitive changes in learning to teach. American Educational Research Journal, 26, 160-189.

Jennings, N.E. (1996). Interpreting Policy Issues in Real Classrooms. New York, NY: Teachers College Press.

Jennings, N.E. & Grant, S.G. (1994). Teaching reading: Teachers learn from policy. In C.K. Kinzer & D.J. Leu (Eds.), Multidimensional Aspects of Literacy Research, Theory, and Practice: Forty Third Yearbook of the National Reading Council. Chicago, IL: National Reading Conference.

Kamii, C. & Kamii, M. (1990). Why achievement testing should stop. In Kamii, C. (Ed.), Achievement Testing in the Early Grades: The Games Grown-Ups Play (pp.15-38). Washington, D.C: National Association for the Education of Young Children.

Khatti, N., Kane, A., Reeve, A. (1995). How performance assessments affect teaching and learning. Educational Leadership, 53, 80-83.

Koretz, D., Stecher, B., Klein, S., McCaffrey, D. (1994). The Vermont Portfolio Assessment Program: Findings and Implications. Educational Measurement: Issues and Practices, 13, 5-16.

Leslie, L. & Caldwell, J. (1995). Qualitative Reading Inventory - II. New York, NY: Harper Collins.

Levin, H.M. & Hopfenberg, W.S. (1991). Don't remediate: Accelerate!  
Principal, 70, 11-13.

Lincoln, Y. & Guba, E. (1985). Naturalistic Inquiry. Newbury Park, CA: Sage.

Lipson, M., Mosenthal, J., Daniels, P., Jiron, H. (under review). The meaning and use of process writing in the classrooms of teachers with different orientations to teaching and learning.

Little, J. (1987). Teachers as colleagues. In V. Richardson (Ed.), Educators' handbook: A research perspective (pp. 491-518). New York: Longman.

Lortie, D.C.(1975). Schoolteacher: A Sociological Study. Chicago: University of Chicago Press.

MacGinitie, W.H., MacGinitie, R.K. (1989). Gates-MacGinitie Reading Tests (3<sup>rd</sup> Ed.). Chicago: Riverside Publishing Co.

Madden, N.A., Slavin, R.E., Wasik, B.A., & Dolan, L.J. (1997). Reading, writing and language arts in Success for All. In Stahl S. A. & Hayes D. A. (Eds.), Instructional Models in Reading. Mahwah, N.J: Lawrence Erlbaum Associates.

Malen, B. & Knapp, M. (1994). Analyses of policy from multiple perspectives: A depiction and appraisal. Paper presented at the American Educational Research Association, New Orleans, LA.

Mathison, S. (1991). Implementing curricular change through state-mandated testing: Ethical issues. Journal of Curriculum and Supervision, 6, 201-212.

McLaughlin, M.W. (1987). Learning from experience: Lessons from policy implementation. Educational Evaluation and Policy Analysis, 9, 171-178.

Madeus, G.F. (1988). The distortion of teaching and testing: High stakes testing and instruction. Peabody Journal of Education, 65, 29-46.

Mehrens, W. (1998, April). Consequences of assessment: What is the evidence? Paper presented at the meeting of the American Educational Research Association, San Diego, CA.

Miles, M. & Huberman, A. (1994). Qualitative data analysis. Thousand Oaks, CA: Sage Publications.

Munby, H.(1984). A qualitative approach to the study of a teacher's beliefs. Journal of Research in Science Teaching, 21, 27-38.

Nelson, B.S. (1999). Building New Knowledge by Thinking: How Administrators Can Learn What They Need to Know about Mathematics. Newton, MA: Center for the Development of Teaching, Educational Development Center, Inc.

Nobel, A.J. & Smith, M.L. (1994). Old and new beliefs about measurement-driven reform: "Build it and they will come." Educational Policy, 8, 111-136.

Osborn, M., Broadfoot, P., Abbot, D., Croll, P., Pollard, A. (1992). The impact of current changes in English primary schools on teacher professionalism. Teachers College Record, 94, 138-151.

Pajares, M.F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. Review of Educational Research, 62, 307-332.

Putnam, R.T. (1992). Teaching the "hows" of mathematics for everyday life: A case study of a fifth-grade teacher. The Elementary School Journal, 93, 163-177.

Resnick, L. & Resnick, D. (1991). Assessing the thinking curriculum: New tools for educational reform. In B. Gifford & M. O'Connor (Eds), Changing assessments: Alternative views of aptitude, achievement and instruction (pp.37-75). Boston: Kluwer.

Richardson, V. (1990). Significant and worthwhile change in teaching practice. Educational Researcher, 19, 10-18.

Richardson, V., Anders, P., Tidwell, D., & Lloyd, C. (1991). The relationship between teachers' beliefs and practices in reading comprehension instruction. American Educational Research Journal, 28, 559-586.

Richardson, V. & Anders, P. (1994). The study of teacher change. In V. Richardson (Ed.), Teacher change and the staff development process: A case in reading instruction. (pp. 159 - 180). New York: Teachers' College Press.

Richardson, V. & Placier, P. (in press). Teacher change. In V. Richardson (Ed.), Handbook of Research on Teaching (Fourth Edition). Washington D.C: American Educational Research Association.

Rosenholtz, S.J. (1989). Teachers' Workplace: The Social Organization of Schools. New York: Longman.

Rosenholtz, S.J., Bassler, O., & Hoover-Dempsey, K. (1986). Organizational conditions for learning. Teaching and Teacher Education, 2, 91-104.

Rowan, B. (1995). The organizational design of schools. In S. Bacharach & B. Mundell (Eds.), Images of Schools (pp. 11-41). Thousand Oaks, CA: Corwin Press.

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

Shepard, L.A. (1995). Using assessment to improve student learning. Educational Leadership, 52, 38-43.

Shepard, L.A. (2000a). The role of classroom assessment in teaching and learning (CSE Technical Report 517). Los Angeles: National Center for Research on Evaluation, Standards and Student Testing.

Shepard, L.A. (2000b). The role of assessment in a learning culture. Educational Researcher, 29, 4-14.

Shepard, L. A. & Dougherty, K.C. (1991, April). Effects of high-stakes testing on instruction. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

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

Sheingold, K., Heller, J., Paulukonis, S. (1994, April). Actively seeking evidence: Shifts in teachers' thinking and practice through assessment development. A paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Smith, M.L. (1991). Put to the test: The effects of external testing on teachers. Educational Researcher, 20, 8-11.

Smith, M.L. & Shepard, L.A. (1988). Kindergarten readiness and retention: A qualitative study of teachers' beliefs and practices. American Educational Research Journal, 25, 307-333.

Smylie, M. (1988). The enhancement function of staff development: Organizational and psychological antecedents to individual teacher change. American Educational Research Journal, 25, 1-30.

Smylie, M. & Conyers, J.G. (1991). Changing conceptions of teaching influence the future of staff development. Journal of staff development, 12, 12-16.

Snow, C.E., Burns, M.S., Griffen, P. (1998). Preventing Reading Difficulties in Young Children. Washington, D.C: 1998.

Spillane, J.P. (1996). School districts matter: Local educational authorities and state instructional policy. Educational Policy, 10, 63-87.

Spillane, J.P. & Jennings, N.E. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. Teachers College Record, 98, 449-481.

Spillane, J.P. & Thompson, C.L. (1997). Reconstructing conceptions of local capacity: The local education agency's capacity for ambitious instructional reform. Educational Evaluation and Policy Analysis, 19, 185-203.

Spillane, J.P. & Zeuli, J.S. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. Educational Evaluation and Policy Analysis, 21, 1-27.

Standerford, N.S. (1997). Reforming reading instruction on multiple levels: Interrelations and disconnections across the state, district, and classroom levels. Educational Policy, 11, 58-91.

Stephens, D., Pearson, P.D., Gilrane, C., Roe, M., Stallman, S.C., Shelton, J., Weinzierl, J., Rodriguez, A., & Commeyras, M. (1995). Assessment and decision making in schools: A cross-site analysis. Reading Research Quarterly, 30, 478-499.

Tobin, K.L. (1990). Changing metaphors and beliefs: A master switch for teaching? Theory into Practice, 29, 122-127.

Torrance, H. (1993). Combining measurement-driven instruction with authentic assessment: Some initial observations of national assessment in England and Wales. Educational Evaluation and Policy Analysis, 15, 81-90.

Valdez, A. (1992). Changes in teachers' beliefs, understandings and practices concerning reading comprehension through the use of practical arguments: A follow-up study. Unpublished doctoral dissertation, University of Arizona, Tucson.

Valencia, S. & Wixson, K. (2000). Policy oriented research on literacy standards and assessment. In Kamil, M., Pearson, P., & Mosenthal, P. (Eds.), Handbook of Reading Research (3rd edition). New York: Longman.

Wixson, K., Valencia, S. & Lipson, M. (1994). Issues in literacy assessment: Facing the realities of internal and external assessment. Journal of Reading Behavior (26), 315-337.

## **Appendix A**

### **Teacher Think Aloud Immediately Following the Second Grade Test**

Teachers will be given a half day of release time to test their children and meet with me for the think aloud. Teachers will select several students, preferably with a range of reading abilities, for administration of the test. Teachers will introduce me to their students individually. I will explain to each student that I am interested in how the test that they are taking helps teachers teach students to read and then thank them for letting us tape record their session. I will not ask the students to identify themselves on the tape. After the student has read the passage orally and answered the comprehension questions (if the district has selected that option), the student will leave the testing area and the teacher will do a think aloud as she scores the student's oral reading.

1. So, tell me what you are thinking as you score the student on each of the rubric components. (accuracy, rate, expression and possibly comprehension)
2. Do you find this child easy to rate? Why or why not?
3. Did you notice anything else about the student's reading?
4. Did the test give you any new information about this child?
5. How do you plan to use the results of this test?

Prompt: Will the test impact the instruction that you give to this child? Why or why not?

6. What kind of instruction will you plan for this child?

## **Appendix B**

### **Semi-Structured Interview Protocol for Teachers**

I will mail teachers the Theoretical Reading Orientation Survey (Lipson, In Press) prior to interviewing them. Ideally, the teacher interview will occur sometime prior to my observations of their administration of the Second Grade Test although this is subject to school and teacher timelines.

1. Tell me about your job.

How many kids are in your class?

How long have you been teaching?

How long have you been at this school?

How long have you been teaching second grade?

What are the things you find interesting? Challenging? Take up most of your time?

2. Tell me a little bit about how reading is taught in your classroom.

What is your philosophy about reading instruction?

From your perspective, what does good reading instruction look like?

What do you think second graders need to learn in reading?

How is your reading instruction organized?

What materials do you use? What do you think about them?

3. How is the implementation of the state second grade reading test going in your school?

What do you know about it? What do you think is the purpose of the test? What do you think about the test?

How did your district decide which test to use?

How did you become informed about the test?

What is going on at the district or school level in relation to the test?

How were the results of the test used last year? By you? By the school? By the district? How do you plan to use them this year?

What impact do you think the test is having having? (Prompt: administrators, children, parents)

4. What is going on in the school that influences your decisions about reading?

Prompt for: School reading policies, organization, other teachers, parents, students

5. What is going on in at the state or district level (central office) that influences your decisions about reading?

WASL/ITBS?

District level policy or decisions about reading?

Change in personnel?

Have you been involved in district policy decisions about reading?

6. Are there other factors in the district that influence reading instruction?

Prompt for: Demographics, parent pressure, teacher turnover?

7. Have you taken any reading staff development in the past two years?

What was it? Who lead it? Why did you take it? Follow-up? What did you think about it?

If offered by the school or district, was the staff development voluntary or mandated?

Were you compensated?

Any other language arts staff development?

6. Anything else you think I should know about reading or language arts instruction in your classroom?

## **Appendix C**

### **Teacher Interview Following Classroom Observation**

1. Was today a pretty typical day in terms of reading instruction?
2. This set of questions will be about various lessons that take place during the reading/language arts period and what the teacher's thinking was behind each one. Why did she chose that activity? Did it work the way she wanted it to? What does she think that the kids learned?
3. This set of questions will be about the students that I observed being tested. What kind of reading work is (child) doing? (Books being read, reading instructional focus, reading tasks) What kind of progress is (child) making?

Thanks a lot -- I appreciate being able to visit in your classroom today.

## Appendix D

### Semi-structured interview -- District Personnel

1. Tell me about your job.

What do you do?

What are the things you find interesting? Challenging? Take up most of your time?

How long have you been the (job) here?

What did you do before this? (Prompt for subject matter and teaching background)

2. Tell me a little bit about how reading is taught at the elementary level in this district.

Do you have any specific policies or practices, either formal or informal, that are focused on reading? (Probe for district outcomes or standards, tests, required text books, philosophies)

How do the EALRs fit in here?

What has been the district response to the EALRs?

Have you done a recent adoption? If yes, what was the process for doing this?

Why did you do an adoption?

Describe how you think reading instruction is handled in the schools in your district.

Describe your philosophy of reading instruction.

From your perspective, what does good reading instruction look like?

Do you see any difference between early grades and intermediate grades?

3. Have you done any reading staff development as a district in the past two years?

What was it? Who lead it? What was the targeted audience? Where did the idea come from? Follow-up? How do you think it worked?

Was the staff development voluntary or mandated?

Were teachers compensated?

Have you offered any other kind of language arts staff development?

4. How is the implementation of the state second grade reading test going in your district?

What do you know about it? What do you think the purpose of the test is?

How did your district decide which test to use?

What is going on at the district level related to the test?

How did teachers become informed about the test?

What do teachers and principals think about the test?

How were the results of the test used? (District, schools, teachers)

What impact do you think it's having? (Prompt: Teachers, children, parents)

5. What other factors do you think are influencing the instruction of reading at this time?

Other state policies? (prompt for WASL, ITBS, grade-level literacy goals and plan)

Changing demographics? Growth? Parents?

Turn-over in teaching force?

6. Anything else you want to tell me about reading or language arts instruction in this district?

Thanks a lot!

## **Appendix E**

### **Semi-structured interview - Principals**

1. Tell me about your job.

What do you do?

What are the things you find interesting? Challenging? Take up most of your time?

How long have you been the principal here?

What did you do before this? (Prompt for subject matter and teaching background)

2. Tell me a little bit about how reading is taught in this school.

Are there any specific policies or practices, either formal or informal, that are focused on reading?

Describe your philosophy about reading instruction.

From your perspective, what does good reading instruction look like?

Do you see any difference between primary and intermediate reading?

What materials do the teachers use?

3. What is going on at the state level that influences what you pay attention to?

(prompt for WASL, ITBS)

4. What is going on in at the district level (central office) that influences what teachers are doing about reading?

District level policy or decisions about reading?

Change in personnel?

Have you or any of your staff been involved in district policy decisions about reading?

5. Are there other factors in the district that influence reading instruction?

Prompt for: Demographics, parent pressure, teacher turnover?

6. What kinds of reading staff development has been available to your staff in the past two years?

What was it? Who lead it? What was the targeted audience? Where did the idea come from? Follow-up? How did teachers respond to it?

Was the staff development voluntary or mandated?

Were teachers compensated?

Any other language arts staff development?

Have teachers participated in state initiated or commercial opportunities to learn about language arts?

What are some of the workshops they've attended?

How have they responded to these?

7. How is the implementation of the state second grade reading test going in your school?

What do you know about it? What do you think the purpose of the test is?

How did your district decide which test to use?

What is going on at the district and school level related to the test?

How did teachers become informed about the test?

What do teachers think about the test?

How are the results of the test used? (district, school, teachers)

What impact do you think it's having? (Prompt: Teachers, children, parents)

6. Anything else you think I should know about reading or language arts instruction in this school ?

Thanks a lot!

**Appendix F**  
**Coding Scheme**

District

Size (Large/Small)

Governance

School-Centered Decision-Making

District Decision-Making

Communication

Professional Development

Literacy

Assessment

DRA/QRI

Other

Central Office Curriculum Staff

Large/Small

Backgrounds

Job Description

Experience

Classroom

Administrative

Knowledge/Beliefs

Reading

Assessment

Perceptions DRA/QRI

Purpose

Positive

Concerns

Reading Instruction

Curriculum

Adoptions

Frameworks

Changes

Assessment

State Mandates

WASL

ITBS

District Mandates

CBA

Gates

STAR

Changes

Stakes

Second-Grade Assessment – DRA/QRI

Use

Development of teacher knowledge

Accountability

Articulation

School

Background

High/low scoring

Big/Small

Evidence of EALRs

Collegiality

Professional Development

Literacy

Assessment

DRA/QRI

Other

Other

Reading Instruction

Organization

Cross-grade grouping for reading

In-classroom instruction

Use of IA's, tutors, volunteers

Materials

Basal series

Trade-books

Accelerated Reader

Second-grade assessment

Use

Resource distribution

Retention

Teacher Accountability

Grouping

School-level data

Development of teacher knowledge

Logistics of administration

Expansion of test

Administrators (Principal, Assistant Principal, Literacy Specialist)

Background

Job Description

Experience

Beliefs about teaching and learning

Literacy

Subject Knowledge

Beliefs

Perceptions of second-grade assessment

Purpose

Concerns

Positives

Teachers

Years in education

Beginning Teacher (1-3 years)

Mid Career Teacher (8-12 years)

Veteran (30-35 years)

Level of education

BA

5<sup>th</sup> year credential

MA

Professional Development

Beliefs

Teaching and Learning

Reading Development

Reading Instruction

Assessment

Practice

Materials

Basals

Workbooks

Trade books

Leveled trade books

Accelerated Reader

Instructional organization

Whole class

Ability grouped within class

Ability grouped across classes

Individualized instruction with leveled trade books

Content and strategies for instruction

Phonics

Comprehension

Fluency

Vocabulary

Enjoyment

Influences on instruction

State

District

School

Other

Assessments

District-mandated

Teacher-selected

Basal End-of-Chapter

Buck Co.

Running Records

Anecdotal Records

AR

Other Assessments

### Second-grade Assessment

#### Perceptions

State purpose

Other

#### Concerns

Time

Levels

Developmental

Standardization

#### Student (pseudonym) Assessment

Administration

Interpretation

Scoring

Instruction

End-of-year

#### Use Statements

Add to teacher's understanding

Dimensions

One-on-one

Other

Parent conferences

Show growth

School Board

Placement

Show standards

Instructional episode

Verification

Modifications

Administration

Scoring

Surprised

Retelling

High fluency, low comprehension

Accuracy

Overall score

## Vita

**Nancy Alison Place**  
**2634 N.E. 82nd St.**  
**Seattle, WA, 98115**  
**(Home) 206-527-0380 (Office) 206-543-6636**

### Education:

Ph.D., Curriculum and Instruction (Literacy, Assessment, Teacher Development)  
 University of Washington, Seattle, 2000.  
 Reading Recovery Teacher Certification. Seattle Pacific University, Seattle, 1994  
 M.A., Human Development. Pacific Oaks College, Pasadena, CA, 1977.  
 Early Childhood and Elementary Teaching Credentials Pacific Oaks College, 1975.  
 B.A., Government. Pomona College, Claremont, CA., 1970.

### Professional Experience:

2000	Teaching Associate and Research Coordinator, University of WA
1999-2000	Research Assistant, Transitions into Teaching: A Longitudinal Study of Beginning Language Arts Teachers, Pamela Grossman and Sheila Valencia, Principal Investigators
1986 -99	Staff Development and Curriculum Specialist, Bellevue Public Schools, Bellevue WA.
1994 - 97	Reading Recovery Teacher, Phantom Lake Elementary School, Bellevue Public Schools, Bellevue, WA
1988 - 99	Adjunct Faculty, Seattle Pacific University, Seattle, WA
1988 - 93	Adjunct Faculty, Pacific Oaks College, Bellevue, WA
1985-86	Teacher, grades 4,5,6, Alternative Program, Skyline School, Solana Beach School District, Solana Beach, CA.
1984-85	Substitute Teacher, Escondido and Solana Beach School Districts, CA
1984 -86	Adjunct Faculty, Pacific Oaks College, San Diego, CA.
1977-84	College Faculty,/Head Teacher, Pacific Oaks College and Children's School, Pasadena, CA
1975-77	Primary Teacher and College Instructor, Center for Early Education, Los Angeles, CA
1975	Home Teacher, Deaf Blind Infants Home Management Program, Charles Drew Postgraduate Medical School, Los Angeles, CA

- 1974 Student Teacher, Orthopedically handicapped classroom, Salvin School, Los Angeles Unified School District, Los Angeles, CA
- 1974 Mathematics Consultant, Rough Rock Demonstration School, Chinle, AZ
- 1973-74 Head Teacher, Cortez Head Start, Cortez, CO
- 1971-73 Teacher, Grades 4 and 5, Rough Rock Demonstration School, Chinle, AZ
- 1970-71 Language Training Staff Member, Japan Overseas Cooperation Volunteers (JOCV), Tokyo, Japan
- 1970 Teacher, English and International Relations, Mindanao State University, Marawi City, Philippines
- 1968 Researcher, Chicago Research Associates, Juvenile Court System, New York, NY

**Honors, Exchanges, Selected National or Regional Presentations:**

Presentations at American Educational Research Association, Washington Educational Research Association, "Every Child a Reader" Conference of the Office of Superintendent of Public Instruction (Washington state), International Reading Association, Washington Commission on Student Learning, National Reading Conference, International Early Childhood Creative Arts Conference, National Council for Social Studies, National Association for the Education of Young Children, California Association for the Education of Young Children, Educational Service District 121 (Washington State), Committee for Children, IX Annual Piagetian Theory and the Helping Professions Conference, North Dakota Study Group on Evaluation. Group facilitation for the San Diego National Geographic Summer Institute, the First Pacific Rim Conference, Vancouver B.C., National Conference of Teachers of English.

1992 Celebrate Literacy Award, given by the Washington Organization of Reading Development for exemplary work in the field of literacy.

1984 Outstanding Faculty Award, Pacific Oaks College.

1982-1983 Faculty Exchange, The Little School, Bellevue, WA.

1980 Faculty Exchange, Bank Street College, New York, NY.

### Notable Projects

- 1997-1998 Project leader, *Early Childhood Literacy Curriculum*, Bellevue Public Schools, Bellevue WA
- July 1994 Elementary Language Arts Room Leader, New Standards Portfolio Project, Indian Wells, CA
- 1993-1994 Team Leader, New Standards Portfolio Project Development
- 1990-1996 Project leader, Language Arts Assessment Project, Bellevue Public Schools
- 1989-1991 Co-Chairperson, Language Arts Revision Committee, Bellevue Public Schools
- 1988 Editor and Project Leader, *Passport to the Pacific Rim*, a third grade social studies curriculum written by Margit McQuire and Rick Moulden
- 1987 Editor and Project Leader, *Bellevue: A Changing Community*, a second grade social studies curriculum written by Margit McQuire and Rick Moulden, Bellevue Public Schools

### Publications

Grossman, P., Valencia, S., Evans, K., Thompson, C., Martin, S., Place, N. Transitions into teaching: Learning to teach writing in teacher education and beyond (in press). Journal of Literacy Research.

Place, N.A. & English, M.(1998). Examining growth over time. In Valencia, S.W. Literacy Portfolios in Action. Fort Worth: Harcourt Brace.

Place, N.A. and Burlingame, C. (1997). Beginning the Day in a Kindergarten Classroom. Video. Bellevue, WA: Bellevue Public Schools.

Silver, B., LaTurner, G., Merlino M., & Place, N. (1994). Jacob Lawrence and Gwendolyn Knight: Artists portraying the American scene. Teacher guide and CD Rom. Bellevue, WA: Bellevue Art Museum.

Valencia, S.W. & Place, N.A. (1994). Portfolios: A process for enhancing teaching and learning. The Reading Teacher, 47: 666-669.

Valencia, S.W. & Place, N.A. (1994). Literacy portfolios for teaching, learning and accountability: The Bellevue literacy assessment project. In Valencia, S.W., Hiebert,

E.H., & Afflerbach, P., (Eds.) Authentic Reading Assessment: Practices and Possibilities. Newark, DE: International Reading Association.

Silver, B., LaTurner, G., Merlino M. & Place, N. (1993). Washington: Voices in Contemporary Sculpture. Teacher Guide. Bellevue, WA: Bellevue Art Museum.

Valencia, S.W. & Place, N.A. (1992). The Bellevue Language Arts Assessment Project. The Notebook. Seattle, WA: College of Education, University of Washington.

Place, N.A. (1991). Assessment as dialogue: Involving students in the assessment process. In Overby, L.Y. (Ed.) Early Childhood Creative Arts: Proceedings of the International Childhood Creative Arts Conference. Reston, VA: American Alliance for Health, Physical Education, Recreation and Dance.

Audap, S., Johnson, D., Jervis, K., & Place, N. (1981). Advice to new teachers: Part III. Insights in Open Education. Grand Forks, ND: University of North Dakota.

Jervis, K. & Place, N. (1980). Teachers learn from teachers. Children's Thinking Newsletter. Cambridge, MA.

Audap, S. & Place, N. (1980). Concrete operational thinking in the classroom. Proceedings from IX Annual Piagetian Theory and the Helping Professions Conference. Los Angeles, CA: University of Southern California.

Place, N. (1978). The reading process observed: Description and analysis. In Jones, E. (Ed.) Readings. Pasadena, CA: National Consortium for Children and Families. Also (1988) in Jones, E.(Ed.), Reading, Writing and Talking with Four, Five and Six Year Olds. Pasadena, CA: Pacific Oaks College.

### **Professional Associations**

International Reading Association  
 National Reading Conference  
 National Council of Teachers of English  
 The Prospect Institute, Bennington, VT.  
 The North Dakota Study Group on Evaluation