The Nature of Data Sources That Inform Decision-Making in Reading by Experienced Second-Grade Teachers

Margaret Mary Dinan Davis

Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/urbanstudies_etds

Part of the Educational Assessment, Evaluation, and Research Commons, Language and Literacy Education Commons, and the Teacher Education and Professional Development Commons

Recommended Citation
Davis, Margaret M. "The Nature of Data Sources That Inform Decision-Making in Reading by Experienced Second-Grade Teachers" (1995). Doctor of Philosophy (PhD), dissertation, , Old Dominion University, DOI: 10.25777/b6z1-ke96
https://digitalcommons.odu.edu/urbanstudies_etds/3

This Dissertation is brought to you for free and open access by the School of Public Service at ODU Digital Commons. It has been accepted for inclusion in Theses and Dissertations in Urban Studies by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
THE NATURE OF DATA SOURCES
THAT INFORM DECISION MAKING IN READING
BY EXPERIENCED SECOND GRADE TEACHERS

by
Margaret Mary Dinan Davis
B.A., College of William and Mary, 1964
M.A., College of William and Mary, 1975
C.A.S., Old Dominion University, 1986

A Dissertation submitted to
the Faculty of
Old Dominion University in Partial
Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

URBAN STUDIES
OLD DOMINION UNIVERSITY
January, 1995

Approved by:
Jack Robinson, Ph.D.
Dissertation Chair
Raymond Morgan, Ed.D.
Member

Jane M. Hager, Ph.D.
Concentration Area Director
Margie S. Tully, Ph.D.
Member

Stephen G. Greiner, Ed.D.
Dean, Darden College of Education

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Abstract
THE NATURE OF DATA SOURCES USED TO INFORM DECISION MAKING IN READING BY EXPERIENCED SECOND GRADE TEACHERS
Margaret Mary Dinan Davis
Old Dominion University, 1994
Director: Dr. Jack Robinson

This qualitative research was designed to describe the nature of data sources that inform decision making in reading by experienced second grade teachers. Data was collected on seven subjects from multiple sources: seven, successive interviews, think alouds using videotaped lessons, classroom observations, and inspection of documents such as grade books, student work samples, report cards, and reading tests. Data analysis was accomplished by transcribing all data into a qualitative data base (Padilla, 1991). Text chunks were tagged and filtered by data source. Categories such as oral language and comparison of data were added as they emerged during analysis. The most important theme to emerge in this study was that of teacher change. While the initial purpose was to describe the assessment data base of seven exemplary teachers, they were all found to be in a state of transition from reliance on basal methodology and comparative assessments to a reliance on whole language methodology and authentic assessment practices. Therefore, their data sources were significantly affected by this transition. The results present a significant contribution by describing the data sources and their use for decision making in reading by experienced teachers in a state of transition. Surveys (Barry, 1992; Coulter, 1992) suggest these subjects are representative of a majority of experienced teachers in transition. Therefore, the results may be generalized to those experienced teachers who are not early adopters of innovation. The results are contrasted with the literature on novices and
recommendations are presented for preservice education and mentorship programs for beginning teachers. The results are then compared to recommended best practices in authentic assessment and recommendations are made for inservice to facilitate teacher change. The findings indicate that teachers in transition must first acquire a knowledge base of reading as a constructive literacy process. Through an understanding of this cognitive activity, they will be able to confront their own belief systems and make meaningful changes in their daily classroom practices. In addition, acquiring the language of reading as an interactive process will enable teachers to articulate better their own intuitive theories about children and learning.
Acknowledgments

I want to acknowledge my deepest appreciation to several people whose constant support has enabled me to wind my way through this journey of discovery and learning.

First of all, I thank my deceased aunt and guardian, Bessie M. Dinan, who believed in me first and inspired me to overcome a poverty of opportunity and spirit.

I acknowledge my gratitude to Elaine Scott Everett, my high school English teacher and life long friend, who gave me the tools to write, and a vision to search for meaning.

I thank my husband, who has warmed and blessed each day.

I thank all my former first graders who taught me the magic of emergent literacy, and left me with many more questions than answers.

I give special thanks to my mentor and former principal, Jean Frances Lowe, who has challenged me to climb many mountains just because they were there.

I thank the seven teachers who participated in this study and all of my professional colleagues who have shared their beliefs, stories, and students with me. It is from this wealth of "kid watching" for the last 25 years that the research questions for this study were framed.

I thank Paula who has always, always been there.

I thank the members of my dissertation committee, Dr. Jack Robinson, Dr. Raymond Morgan, and Dr. Margie Gaines Tully, who have listened and directed me during these three years.

I thank the Virginia Educational Research Association.
for their interest in, and financial support of this project.

Finally, I thank all those who bring music to my life. This gift has carried me through, and added much joy and beauty.
# TABLE OF CONTENTS

Acknowledgments ...................................... ii  
List of Tables ......................................... vii  
List of Graphs ......................................... viii  

CHAPTER 1 INTRODUCTION ............................... 1  
  Statement of Issues ................................... 1  
  Statement of the Problem ............................ 12  
  Purpose of the Study ................................ 13  
  Research Questions ................................... 15  
  Definition of Terms ................................... 15  
  Assumptions ......................................... 17  
  Significance ......................................... 18  
  Limitations ......................................... 20  

Chapter 2 RELATED LITERATURE ....................... 23  
  Overview ............................................ 23  
  Changing Paradigm of Reading and Assessment ... 23  
  Teacher Cognition and Decision Making .......... 25  
  Teacher Judgment .................................... 26  
  Novices and Experts ................................ 29  
  Teacher Change ...................................... 31  
  Data Studies ........................................ 32  
  Conclusion .......................................... 36  

Chapter 3 PROCEDURES ................................. 38  
  Introduction ........................................ 38  
  Research Method ..................................... 38
References ........................................... 276

Appendices .............................................
A. Interview Guide ............................... 284
B. Examples of Interview Data ............... 294
C. Examples of Notes and Documents Provided ... 306
LIST OF TABLES

Table 1  Summary of Percentages for each Data Source 54
Table 2  Data Frequencies for Roberta ................. 56
Table 3  Data Frequencies for Blaine ................. 69
Table 4  Data Frequencies for Amy ................ 89
Table 5  Data Frequencies for Betty ................. 115
Table 6  Data Frequencies for Scotty ................. 140
Table 7  Data Frequencies for Stacy ................. 172
Table 8  Data Frequencies for Kitty ................. 205
Table 9  Reliance upon Word Recognition Data ....... 228
Table 10 Reliance upon Comparison of Data ......... 233
Table 11 Reliance upon Observational Data ........ 237
Table 12 Reliance upon Interactive Data .......... 240
Table 13 Reliance upon Basal Test Data .......... 242
Table 14 Reliance upon Behavior & Word Habits Data 244
Table 15 Reliance upon Oral Response Data ........ 246
Table 16 Reliance upon Oral Language Data ........ 249
LIST OF GRAPHS

Figure 1 Aggregates of Data Sources ............... 262
CHAPTER 1
INTRODUCTION TO THE STUDY

The issues, the problem, and the purpose of the study are presented in this chapter. In addition, generally focused research questions are included as well as definitions of terms. Finally, the assumptions, significance of the study (including its relevance in urban settings), and limitations are discussed.

Statement of Issues

At the present time, the nature of reading instruction is in a state of transition (Tchudi, 1991; Perrone, 1991). The task of reading is being redefined as an integrated language activity rather than a succession of subskills. This follows several movements in the last three decades where the understanding of how learning takes place has shifted back and forth. The Open Classroom movement of the 1970's was accompanied by a new reading paradigm of language experience (Stauffer, 1970) and psycholinguistic advances (Smith, 1971) that proposed that acquiring literacy followed the same cognitive development as learning to talk. Indeed, Stauffer made the point that "reading is talk written down." The 1980's included a return to emphasis upon accountability and development of skills. These were most often measured by standardized tests. Resnick and Resnick (1992) point out that each wave of reform has been accompanied by a redefinition of testing and assessment.

Cognitive researchers such as Resnick (1987) have been responsible for the major paradigm shift in how learning is perceived to take place. Reading is no longer perceived to be an accumulation of skills, where knowledge is added up somehow until teachers fill the beaker and the student reaches a criterion score. Rather,
reading is conceptualized as a constructivist activity wherein the learner interacts with the text within the reading context to construct meaning. This process is mediated by the learner's prior knowledge, beliefs and interests, as well as the ability to monitor understanding and adjust strategies during the reading process.

Teaching, therefore, is being redefined as the process of facilitating this engagement of students. Assessment then, must accommodate this emerging model by allowing the teacher to receive, interpret, and value data regarding the reader's ability to construct meaning, and to provide additional experiences where appropriate (Goodman, 1989). Assessment tasks must be more authentic or closer to the instructional task in order to achieve a more valid measurement, and to enable the teacher to gain immediate feedback on student needs for modifications (Harp, 1993).

At the same time, standardized testing has received a decreased emphasis for use in instructional planning. Educators have recognized that these data have not provided sufficient information regarding the interactive processes of learners. In addition, test data have often narrowed the curriculum as teachers responded to external mandates to raise scores and thereby focused instruction on tasks that fostered convergent lower level thinking skills.

A concomitant shift has occurred in methodology which translated the view of interactive learners and facilitating teachers into an instructional belief system that values real literature, multiple texts, constructed responses to reading, and rich, meaningful assessment data that give a window into meaning-getting processes (Goodman, 1989). This is what Goodman has dubbed "kid watching"—discovering with children how and what they know and why it is important.

The basal texts that provided teachers with a map of discrete skills to cover, and tests that were accompanied by decision making
parameters in the form of criterion scores, have given way to reading-writing classrooms. Here the classroom practitioner must make decisions regarding materials and methodology based upon specified learning targets. Collection of assessment data becomes a function of external mandates, teacher beliefs and knowledge, and the nature of the decision to be made. However, teachers' preservice training may have omitted the basic elements of "assessment literacy" (Stiggins, 1992) and they may be unable to perform many of these responsibilities.

Not enough is known about the collection, valuing, and use of data in the language arts classroom. Furthermore, while a great deal is known about beginning teachers from novice studies (Berliner, 1984) and planning studies (as reviewed by Clark and Peterson, 1986) less is known about the use of data and the decision making strategies of experienced teachers. Prediction studies revealed that teachers' ability to predict student success on standardized language arts measures is characterized by 60% accuracy (Gaines and Davis, 1990), a .78 correlation (Oliver and Arnold, 1978), and an increase from .55 to .77 as the year progresses (Morine-Dershimer, 1979). Moreover, the latter study reported that prediction accuracy is raised as the outcome measure approximates the instructional task. Gaines and Davis (1990) reported prediction accuracy to be a function of experience. Those teachers with twelve to fourteen years of experience demonstrated greater accuracy.

While these studies imply that predicting student success includes the effective use of student achievement data, the prediction literature does not specify data sources. In a similar fashion, the literature on decision making research appears to focus mainly upon models of decision making. While input data are included in the decision making models of Shavelson (1977, 1981) and Borko et al. (1979), specific data sources were not reported.
Finally, attribution literature appears to overlap with prediction research. While prediction literature focuses upon teacher judgment prior to measurement, attribution research focuses upon teachers' thoughts about achievement after it has been measured as they attribute learning to variables. These include ethnicity (Cooper, Baron and Lowe, 1975), socioeconomic status of students (Rist, 1970), variance of class (Kagan, 1988), and ability or behavior and effort (Cooper and Burger, 1980). Again, specific student achievement data were not included in any of these attribution studies.

A baseline, therefore, for training teachers in collecting and interpreting assessment data is the behavior of those teachers with demonstrated expertise in making accurate judgments about student achievement. Their data sources must be clearly described. Investigating teachers who are judged by others as knowing their students well may shed light on the data sources they use to make their judgments about students. How they use observational data to make decisions during instruction may have significant implications for the design of teacher-training programs (Chittenden, 1991). This training might include emphasis upon the interpersonal dimensions of classroom assessment (Stiggins, Griswold, and Wikelund, 1989). In addition, teachers must understand the increasing reliance upon teacher judgment in interpreting all data available for decision making. This is especially true of informal and observational data (Wang, 1988).

As early as 1986, the assessment community began to document the need for additional research in determining how often informal data such as observational and anecdotal records were being collected (Farr and Carey, 1986). This research agenda also included the validity and reliability of such data, and their use in informing the decision making process. Farr and Carey conclude their treatment of
the state of measurement in reading with the following statement:

We need continued research on the types of decisions made by teachers, curriculum supervisors, and administrators to plan and carry out instruction. We also need to know the kinds of information they require to make those decisions, the format in which the information should be provided, and how this relates to the timing of decision making. Such research most likely will reveal that educators need a wide variety of information for decision making, much broader than that which could be provided by any reading test. Moreover, these information needs probably cannot wait until a test is administered and scored and a set of packaged results are returned to the teacher or administrator. We know that information needs for planning instruction are often immediate. Such immediate needs necessarily rely on informal evaluation carried out as part of ongoing instruction, an area of reading assessment which has already been described as woefully underdeveloped (p. 213).

Finally, in an analysis of assessment decisions and how to inform those decisions, it becomes apparent that assessment currently addresses three distinct purposes: classroom diagnosis, instructional monitoring and revision, and accountability. It is not clear whether the same assessment data can meet these three different purposes. Assessment purposes drive decisions about what information is collected and how the information is used. Assessment data collected for accountability and for reporting to parents is often different
from information collected for determining students' instructional needs. In addition, data collected to serve assessment purposes beyond the classroom, such as for sorting and selecting and for program evaluation or for policy decisions must meet more rigorous psychometric standards for reliability, validity, fairness, and generalizability.

For example, instructional monitoring and revision questions might include: Is this novel appropriate for all students? Did the lessons on brainstorming and mapping help students organize their paragraphs? Are there any students who do not choose recreational reading in their free time and why? Are all students responding orally during comprehension discussion? Which students have limited receptive vocabulary? How successful are students in utilizing context clues to derive meaning? All of these questions require observation, judgment, and collection of different types of data. Other student achievement data already available, such as standardized test data, may assist the teacher, but it cannot substitute for data collected inside the classroom.

Conversely, accountability and program evaluation decisions are less concerned with individual student performance, and more concerned with class, school, district, or state achievement. These questions might include: Is a certain program successful in reducing dropouts? Is the disparity between males and females in math widening in middle schools? Do middle school students enter with reading skills sufficient to profit from instruction in content area texts? What areas of the state have the lowest or highest dropout rates? What variables in the disaggregated data appear to contribute to this pattern? These questions require data that can be obtained efficiently, possess a high degree of reliability and comparability, and can be aggregated across levels or variables to assist in tracking trends, cohorts, or subgroups.
Therefore, data collection in a school system or state can no longer be of one type (formal vs. informal) exclusively (Stiggins, 1994). A teacher who has taught a student successfully for three months and then observes he is not mastering instructional objectives, needs student observational data, student work samples, and perhaps student interview data. This information would not be sufficient, however, for the system superintendent to decide whether to fund an expensive individualized reading remediation program for a third year in the school system.

Teachers collect and react to hundreds of bits of student data on a daily basis. Stiggins and Conklin (1992) include over 400 cues in their descriptions of classroom assessment contexts. Gall (as reported by Stiggins and Conklin) estimates that teachers ask 150 questions per hour. Shavelson and Stern (1981) identified 66 cues used in making instructional decisions, and Clark and Peterson (1986) estimate that these decisions occur every two minutes.

Experienced teachers make hundreds of instructional decisions weekly in their classrooms. While a great deal of attention in the literature is devoted to examining teacher cognition and decision making (Clark and Peterson, 1986; Clark and Yinger, 1979; Shavelson and Stern, 1981), much less effort has been expended in describing the assessment data that inform this decision making process and the factors which shape its collection, valuing, and use in modifying instruction (Madaus and Kellaghan, 1992). Stiggins and Conklin (1992) hypothesize about the lack of investigations:

... the extreme complexity of the classroom assessment environments and issues have served as a barrier that has kept researchers from conducting the needed research (p. viii).

Therefore, it is essential to provide an elaborate, verbal
description of teacher thinking in the area of reading assessment. It is important to explore the types of decisions made by teachers, specialists, and administrators, the types of data required to inform decision making, the most usable form for these data, and information on the sequencing and timing of decision making (Farr and Carey, 1986; Pryor, 1992).

Little is known about the assessment behavior of beginning teachers although their teaching behaviors have been studied in depth (Hollingsworth, 1989). This is most likely due to their preoccupation with managerial issues in the first year(s). Beginning teachers become sensitive to assessment data at the classroom level only after classroom management becomes routine (Hollingsworth, 1989). Initially, they appear overcome by the complexity and simultaneity of the classroom environment. Only after they develop and implement routines that reduce the amount of information they must process, are they able to attend to interactive behavior during instruction. They then begin to process these cues and make attributions about student learning based upon these data. It is not known how long this process takes.

Chittenden (1991) recommends an investigative study of the assessment practices of exemplary teachers in the classroom by all school systems. In this way, initiatives in school districts might build on these practices and extend them. Even more specifically he states, "A useful question to pose to teachers is something on the order of: What are indications to you that the child is making progress as a reader? What does the child do? Not do? When, Where?" (p.27) Stiggins (1989) makes the most persuasive argument for study of master teachers in order to improve teacher education:

Assessment instructors must understand the realities of life in classrooms. All who have not spent time in
public school classrooms or haven't been there recently should go and observe and teach there. This will reveal to them the complexity of the assessment task demands teachers face every day. Through this in-class experience, assessment instructors also can learn from good teachers the basic principles of good teaching. These principles can be applied to the development and presentation of sound assessment instruction also (p. 10).

Dole, Duffy, Roehler and Pearson (1991) describe the teacher as an interactive decision maker during reading. As the lesson progresses, the instructor receives responses from students and modifies existing plans to meet the needs of students through response elaboration. Berliner (1984) reports that teachers make ten decisions per hour that are "instructionally significant." Leinhardt as reported by Bransford and Vye (1989) extends the investigation of the new teacher's ability to process data in the classroom from the management issues to include his or her predictive or attribution skills. As new teachers begin to manage the simultaneity of the classroom, they begin to revise initial estimates of students' abilities and needs. In this way, the teacher is able to predict students' success with given materials or tasks, and build a schema for organizing, assessing, and managing instruction.

Harp (1993) underscores intuition as the most important source of assessment data in whole language classrooms where data collection and assessment decisions are not prescribed by the basal management system. However, he argues that teacher judgment is only a valuable source of decision making data if it is "based on careful observation and knowledge of a child's learning." Therefore, conducting an extensive study of teachers with recognized expertise in observing and interpreting information about students' performance could have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
value in planning teacher training programs. Careful documentation of teachers' data sources and description of how these data are valued in the decision making strategies used by these teachers, will provide grounding for development of alternative assessment models. A rich, verbal description of their strategies and data sources could also lead to improvement in beginning teacher assistance programs. While a teacher's first year is spent on management issues, renewed effort must be devoted to second and third year teachers who are in the process of developing and revising their assessment strategies as they attend to new data sources in their classrooms.

Identifying teachers' assessment data sources might also lead to suggesting whether or how to standardize the contents of portfolio assessments. This is especially important if portfolios are used for accountability.

For example, some districts prescribe that portfolios contain specific samples of student work as well as oral reading inventories, etc. (Harp, 1993). Documenting the reliance of experienced teachers upon data sources for decision making may provide data for reflection by administrators and policy makers who often formulate external constraints in this area. If standardization of portfolio contents occurs, the results may be more efficient if selected components are documented to be rich sources of data for experienced teachers. Inservice issues would then revolve around systematically gathering and accurately interpreting these data. In summary, the goal is to understand how teachers who know a great deal about their students' learning, acquire and use this knowledge. The resulting description may inform various audiences such as policy makers, teachers of preservice teachers, and providers of inservice.

This study, therefore, will attempt to describe the current practices in assessment in the York County Public Schools among experienced second grade teachers who demonstrate expertise in
utilizing assessment data, both formal and informal. Experienced will be defined as having taught more than ten years in the primary grades. Demonstrated expertise will be defined as a judgment by the principal and reading specialist in individual schools. They will be asked to nominate subjects who fit the two requirements stated. Their judgments are considered to be a robust measure for selection because they possess the greatest level of training in reading assessment in the school and because they have ongoing dialogues with individual classroom teachers about assessment issues and the progress of individual students.

For example, principals and reading specialists dialogue with teachers throughout the year about the results of countywide assessments in reading, the results of teacher made tests for thematic studies or novels, results of state mandated standardized testing, documentation of mastery of learning objectives, data to support referrals to special education and gifted committees, and data to support grading and promotion or retention decisions. These discussions allow teachers to interpret aloud data on their students, and to compare them to their intuitive judgments of what students have truly learned. This exposure to teachers, as they think aloud about assessment, allows principals and reading specialists to judge their expertise.

By documenting the process of data gathering in language arts, the valuing of these data, and use by experienced teachers in making a variety of instructional decisions, it will be possible to gain information relating to the following questions: What classroom data are most valuable to teachers? What is the mix of formal to informal measures? How are data valued in instructional decision making? How do data sources vary with the types of decisions teachers make? Are instructional and accountability decisions supported by the same types of data? What provision for error reduction and verification
procedures do teachers use in documenting their decisions? How do teachers use data to justify the attributions they make about students?

Statement of the Problem

Systematic knowledge is lacking about the range and types of decisions made by classroom teachers in reading and the types of data used to inform these decisions. In addition, there is little knowledge about the interaction of experience, teacher thinking and beliefs, decision making strategies, teacher orientation toward a methodology, or external mandates, real or perceived, on the types of data collected, valued, and used (Antonacci as reported by Pryor, 1990; Borko, Flory and Cumbo, 1993; Shavelson and Stern, 1981; Harp, 1993; Stiggins and Bridgeford, 1985).

For example, teachers who use a basal with accompanying tests and decision making parameters may not collect any additional data other than work samples. However, if they teach students with word recognition difficulties, they may collect and value oral reading or running record data. In this case, data collection is driven by student need rather than methodology. Teachers who perceive their students to be highly fluent readers may not value oral reading data, or they may attend to different features such as a student's ability to demonstrate phrasing, intonation and expression. Some teachers report attending to a student's oral language as an estimate of ability. While this appears to fall into the area of teacher cognition and attribution, it does have important implications for data collection. Once this perception of ability estimate is established, subsequent data are compared to it. Thus, a highly verbal child with well-developed critical thinking skills, when measured orally, who also displays severe decoding needs may cause
experienced teachers to question, collect more data, and attempt to resolve the lack of consistency observed. Teachers who report using real literature in their classrooms also report collecting children's responses to literature, usually in writing. Finally, teachers who work in districts that mandate writing folders or portfolios, with numbers and types of writing products specified to be collected, tend to structure their classrooms in order to collect and value these data.

Pryor (1992) points out the paucity of information on how teachers value information for grading, report cards, and reporting to parents. Finally, there is a lack of data regarding the behavior of experienced assessors in reading to inform preservice and inservice programs. Teale, Hiebert, and Chittenden (1987) underscore this need by reporting a discrepancy between knowledge about emergent literacy and assessment practices in the classroom. Although Goodman, Smith, Meredith and Goodman (1987) and Clay (1985; 1982) have chronicled the development of the young child through developmental stages of prereading and prewriting to emergent literacy, less attention has been devoted to the development and implementation of assessments to track this development.

Purpose of the Study

The purpose of the study is to provide descriptive information about the classroom assessment practices of seven experienced second grade teachers, with demonstrated expertise in collecting and interpreting assessment data. In addition, the study will compare the assessment methods of these teachers to a model of literacy which might suggest the types of data that are needed to track the development of literacy in children. Describing teachers' data sources for making judgments about students and comparing these data to a literacy model may lead to identification of discrepancies.
This, in turn, might suggest directions for teacher training.

These data add to the increasing number of case studies in this area. Moreover, the successive interview design of this study provides a substantial increase in the amount of data available on teacher thinking about assessment, and strategies teachers employ to collect and value data to make instructional decisions in reading. This is achieved by conducting successive interviews over the course of several months focused around themes such as grading, classroom organization, formation of instructional groupings, testing, etc. In addition, data sources are not specified in the interviews; therefore, the contamination of data by suggesting or narrowing data choices (as through surveys) is avoided.

For example, surveys typically ask teachers to rate the importance of data sources (such as homework completion or accuracy). Surveys may ask teachers to distribute a number of points among data sources to reflect their thoughts on the relative importance of different data sources in decision making. However, by asking teachers generally focused questions such as "how were your groups formed, and how have they changed?", teachers are encouraged to talk aloud about their professional role as reading educators. Data sources emerge naturally. Once teachers state a data source such as oral responses to comprehension questions, they are probed for additional information with general probes such as "could you tell me more?" or more specifically, "could you explain how their oral responses help you know who understands?"

Stiggins and Conklin (1992) began their investigation of teachers' assessment practices with surveys and then found they must interview and observe in order to capture the actual complexity of the classroom context and accompanying assessment demands. They chose ethnographic methods because they provided the rigors of research to inform practice but also allowed the flexibility needed
to describe this interactive process.

**Research Questions**

Several generally focused questions framed this study. They include:

1. When reflecting orally, what do teachers cite as data sources to inform or confirm their judgments in reading?

2. When reviewing documents (daily work, criterion referenced tests, norm referenced tests, video tapes of instructional/assessment activities, and/or scripted student responses), what verbal comments indicate how these data are valued and used in the decision making process?

3. When teachers are asked to reflect upon their deliberate use of assessment strategies and how these have developed over the years, what trends are evident to suggest the effect of experience upon use of assessment data?

**Definition of Terms**

Several key concepts were either crucial to the study or were operationally defined through the interview process. Each is defined for the purpose of the study below.

**Assessment.** Assessment is the process of gathering or attending to any information about students or their performance for reflection at that moment or at a later time.

**Authentic Assessment.** This method of assessment involves gathering data from tasks that are real, "worthy" (Wiggins, 1989), or closer to the real-life context in which a student may apply knowledge gained.

**Domain Scoring.** This is a technique for scoring writing to reflect emphasis upon a specific area of interest in writing such as organization, or mechanics. This method of scoring is used in grading the Virginia Literacy Passport Test in Writing.
Evaluation. Evaluation is the process of reflecting upon and making judgments about assessment data for the purpose of monitoring student progress or program features or for making policy decisions.

Invented Spelling. This is an instructional strategy for writing. Students are encouraged to spell words as they hear them phonetically.

Journal Writing. This is an instructional strategy utilized in many language based classrooms. Children from all grades are invited to record their thoughts, pictures, and ideas in a notebook. These are neither corrected by the teacher nor graded. There is no intervention by the teacher to correct spelling, punctuation, or grammar in order to respect the personal ideas of the child and his or her attempts to construct a written message.

Language Experience. This is a reading philosophy and methodology based upon the work of Stauffer (1970; 1980). It relies heavily upon using a child's oral language as the basis for developing reading skills. It is based upon the premise that "reading is talk written down."

Mentorship. Mentorships may be both the formal and informal process of pairing an experienced and a new teacher to benefit the newer teacher through the sharing of expertise.

Peer Revision. This is a cooperative learning strategy which may be applied in the area of writing. Two students work together to edit a child's creative writing.

Performance Assessment. This is the gathering of data on students as they perform real tasks such as writing a story, reading orally, performing in a play, etc.

Portfolio. A portfolio is a purposeful collection of student products as well as checklists of performance assessments to document student achievement and growth over time.
**Reading Response.** This is a reading-writing strategy. After children listen to literature or read independently, they are encouraged to record their thoughts and feelings about the story, its characters or any other reactions.

**Sizing-up strategies.** Sizing-up strategies include those strategies teachers employ at the beginning of the school year to gain information about the range of needs, abilities, or achievement in the class and to begin grouping students for instruction.

**The Writing Process.** This is a multi-step methodology for teaching writing that was externally mandated in this school division. The process begins with prewriting and proceeds through first draft, revisions, and final draft. Optional elements of the process include using graphic organizers and maps for planning, brainstorming, peer assistance in revision, teacher-pupil writing conferences, and writing portfolios or writing folders.

**Assumptions**

There were a number of assumptions made regarding teachers, their thinking about students, and their assessment of children's reading. These are listed below:

1. It is assumed that all teachers, to varying degrees, collect data on student achievement in reading.

2. It is assumed that data collected regarding student achievement affect, to varying degrees, teachers' thinking about students and their subsequent decisions regarding planning, grouping, and instruction.

3. It is assumed that there are variables that affect teachers' thinking other than student data. These include teacher beliefs, teacher knowledge, and intuition.

4. It is assumed that teachers can verbalize their thinking about students to include their instructional decisions and the data
sources that inform them.

5. It is assumed that teachers are subject to real or perceived external mandates which may or may not be congruent with their belief system.

6. It is assumed that external mandates affect teacher thinking, data collection, and decision making.

**Significance**

This study occurs concurrently with a major paradigm shift in how learning and, therefore, reading and assessment are being redefined. Methodological changes have appeared as instruction makes use of big books, trade books, multiple texts, or real literature rather than basals, and assessment of learning relies upon data from performance of real tasks or constructed responses. Without the criterion referenced tests and decision making parameters in the form of criterion scores, teachers must now assume the responsibility for collecting appropriate data for a variety of decisions about pupils, valuing those data, and utilizing them in the decision making process. Students must also adjust to this change in measurement by intuiting the values underlying authentic assessments. Recognition of correct answers clearly requires different cognitive processes for students than constructed responses that might involve synthesis, comparison, or application. Teachers must acquire skills in articulating these new data to parents; parents must adjust to information that may tell more but answer very different questions by providing unique information about their own child's learning.

Stiggins (1992) has called for training for teachers in order to equip them for the tasks outlined here. Few colleges include measurement in their preservice education for teachers; those that offer measurement classes often devote substantial portions of instructional time to standardized measurements of learning.
Therefore, teachers with and without experience may lack the basics of assessment literacy.

This study provides a significant contribution to the data on classroom assessment practices. By conducting repetitive interviews it is possible to gain insight into teacher thinking. The rich, verbal description of their responses adds to our knowledge of the data sources they use in their decision making about pupils, the mix of cognitive and noncognitive data, and the strategies employed in valuing data. Because only experienced teachers are subjects, it is possible to view these activities independent of managerial issues which mediate data collection by beginning teachers (Hollingsworth, 1989). The shift in emphasis upon criterion referenced basal tests to more authentic assessment, and the research setting here which is characterized by transition, offer an opportunity to view teacher change behavior and emerging issues that are a function of adaptation and assimilation of new external mandates. Finally, as Stiggins (1989) and others emphasize, preservice education and inservice education for teachers must be informed by a current understanding of the "assessment literacy" of experienced teachers in the field, as well as their behavior in adapting to change. It is these individuals that have the greatest access to influence novice teachers through student teaching and formal or informal mentoring programs.

This research problem is of particular significance in urban settings. Accurate and valid assessment data are essential in providing early intervention and modifications through instructional programs for at-risk students. The implementation of authentic assessment practices is of particular importance in multicultural, urban settings. The data that are produced from performance assessments, anecdotal records, journals, literature responses, and creative writing describe rather than compare students of diverse
populations. The underlying set of beliefs for valuing these data accepts and respects their varying prior knowledge, beliefs, and language. Constructed responses allow greater insight in understanding what students comprehend, and how they learn. This is especially important in planning instruction for multicultural students as well as those with limited English proficiency. Finally, descriptive data allow a closer link to accurate instructional placement for transient students. This need is well documented in urban settings (Ascher, 1993; Ascher, 1990). All teachers must possess assessment expertise needed to collect, interpret, value, and act appropriately on this information.

Limitations

There are several limitations that affect this study. First, interviews were primarily conducted between January and June and might have omitted data about planning and sizing-up assessment activities that are typically conducted by teachers at the beginning of the school year to make grouping and other instructional decisions. Questions in the initial interview focused upon how groups were established in the classroom while questions in the exit interview invited informants to share sizing-up strategies. In this way an attempt was made to overcome this limitation.

Another limitation of an interview study might be the possible discrepancy between teacher reports of their thinking and practices and their actual classroom behavior. Multiple interviews with overlapping themes, classroom observations, and videotaping allowed triangulation of the data to identify such discrepancies.

A related limitation of survey studies might be the leading of informants by specifying data sources. The formation of interview questions guarded against this. In addition, requests for additional
information were framed from a subject's previous responses (i.e., "You said previously you look at ____ during discussion of the text. Would you tell me more about what you look for or how you use this information?").

Another limitation might have been that data gathering and decision making in reading are a function of grade level. Assessment issues in emergent literacy in kindergarten and first grade may require quite different data from the issues regarding students who are expected to read and solve problems, construct a model, or evaluate propaganda in the upper grades. This limitation was addressed by confining the study to second grade and interpreting the findings within that context.

A limitation might have resulted from the fact that the research setting, The York County School Division, was in a state of transition. That is, they continued to require teachers to give a criterion referenced basal test, while concurrently providing inservice on authentic assessment practices and whole language. This limitation was minimized by carefully documenting the setting and nature of external mandates. Moreover, it should be noted that York appears representative on the change continuum among school districts in Virginia, according to a Virginia State Reading Association Survey (Coulter, 1992) conducted at that time. This revealed a majority of Virginia second grade teacher respondents employed both basal and whole language methodology. The seven subjects in the present study reported use of basal and trade books for instructional use. Another 1992 survey (Barry) reported 69% of 206 respondents in a southeastern state use a basal as the primary mode of instruction. Thus, it appears that the subjects in this study may share characteristics with many teachers in the southeast.

Moreover, data on teacher change emerged as function of this context and is a significant outcome of the study. Interview data
collected in the context of a changing paradigm of methodology and assessment, reflect teachers' thoughts, belief systems, and concerns.

A final limitation was that half of one interview with one of the subjects, Kitty, was inaudible for transcription. Transcription from field notes and papers provided and the collection of multiple sources of data lessened the impact of this malfunction of equipment.
CHAPTER 2
RELATED LITERATURE

Overview

This chapter presents a review of literature related to the research questions of this study. Specifically, the changing paradigm of reading and its assessment is examined. A number of related issues are reviewed in the area of teacher cognition. These include teacher thinking, expectations, attributions and beliefs, teacher decision making, teacher planning and teacher change. Aspects of teachers' thoughts that affect the selection, collection, valuing, and comparison of data are reviewed briefly in this chapter. Novice and expert studies are also briefly examined in light of results in the area of processing data and decision making. Finally, studies which have targeted teachers' data sources are considered in greater depth.

Changing Paradigm of Reading and Assessment

The prevailing definition of reading through the seventies and beginning eighties was the summation definition: reading is a summation of discrete subskills taught in careful sequence and becoming progressively more complex. Valencia and Pearson (1987) attributed the popularity of this definition to the twenty year favor of mastery learning based on the prevailing behaviorist theory of learning. Through this model educators had dissected learning into small component parts to be practiced separately. The teacher controlled pacing of new skills introduced and the amount of practice required. Teachers attempted to match materials and method to the needs of the learner. Emergent or young readers in this model were viewed as unable to become expert readers until they had acquired all
the component skills (Wiggins, 1989).

Glazer and Searfoss (1988) perceived a pendulum effect from defining reading as a process early in the century, then as a product, and finally returning to the process orientation. Many events produced a major shift in the conceptualization of reading during the seventies and eighties. Frank Smith (1971) emerged and labeled reading a psycholinguistic guessing game. Building on the concept that language underlies all literacy learning, Ken Goodman called for meaningful authentic texts for emergent readers, while others declared reading an interactive process (Pearson and Spiro, 1980). In his later evaluation text, Goodman (1989) invited teachers to extend the interactive process to assessment by adopting a stance as "kid watchers," trying to discover strategies readers use as they bring prior knowledge to the text and then attempt to construct meaning. The International Reading Association brought national attention to the issue of disparity between progress made in curriculum reform and that made in assessment. This was accomplished by devoting an entire issue of The Reading Teacher to this theme in April of 1987.

The implementation of reading curriculum reform is being accomplished more smoothly than an accompanying assessment change. There are several reasons for this. First of all, sampling or recognition tests are perceived as more manageable by many teachers and administrators, and as more reliable by many policy makers and parents. In addition, policy makers value data that are easily aggregated and compared. The construction of tests that accurately document the interaction of receptive vocabulary, prior knowledge, language, beliefs, and experiences with the surface level of the text and reflect the process of making meaning, presents a myriad of psychometric challenges (Glazer and Searfoss, 1988). Jan Hancock (1992) simplifies the verbose description above by stating that the
Latin root of assessment is *assidere* which means "sit beside." According to Hancock, when one sits beside students, one comes to know them.

**Teacher Cognition and Decision Making**

Teacher cognition is presented as an umbrella under which to discuss briefly all those aspects of teacher thought that relate to the research questions in this study. These include teacher expectation, teacher judgment, teacher beliefs, planning, and decision making.

The ability of teachers to assess students' performance and needs accurately is an essential component of any reading assessment program. Other than structured criterion referenced multiple choice tests with clear parameters for grading and decision making, all data require judgment by the teacher to interpret and value. Teacher expectancy is essential in order for teachers to perform several important tasks: planning, grouping for instruction, selecting students for special programs, and selecting materials and methodology for instruction (Cooksey and Freebody, 1985). In addition, teachers develop schema over time that enable them to reduce the amount of information they must process and enable them to deal more effectively with the simultaneity of the classroom (Doyle, 1979).

Certifying competency in one area and readiness for another is not an exact science in a language based classroom. Even judging whether a given response satisfies a rubric on an authentic assessment requires several mental processes. These include recalling the rubric and exemplars of given levels of responses, accurately recording the students' response, consideration of the context of the response, comparison of the response with the exemplars, and making a decision about the quality and completeness
of this response. Even once this decision is made, an accompanying process of instructional revision begins. Teachers alter their instructional techniques based on students' responses.

Teacher prediction of student achievement begins in kindergarten where teachers assess readiness for reading, and make recommendations for first grade teachers (Kermonian, 1962). In addition, they may make recommendations of students they judge "at-risk" for inclusion in early intervention programs. By intervening as soon as possible with correctly identified students, the greatest gains may be made (Kagan, 1988). The accurate identification of students for at-risk programs impacts the school system's effectiveness in delivering these programs. By under-identifying, delays in intervention result which affect student progress. By over-identifying, the division's efforts are diluted, and effectiveness may be decreased (Gaines and Davis, 1990).

Teacher Judgment

Teacher judgment influences not only which data to collect on students' achievement, but how to document, value, and report those data. It also includes the variables that influence or bias judgment. Phillip Jackson (1968) provides important baseline information from his study of classroom environments. His work concludes that teaching, teacher thinking, and classroom contexts are complex, multifaceted, and impacted by many variables. Some of these are related to achievement. Some are tangential; others are extraneous.

As mentioned, teachers develop schema over time to help them deal effectively with the simultaneity of the classroom. These cognitive schema assist them in categorizing events including behavior, responses, performance indicators, etc., in order to reduce the complexity of the classroom context (Doyle, 1979). In this way,
they may process some information automatically in order to focus on selected data. This automaticity frees up the teacher's cognitive processing energy.

Teacher planning studies such as Yinger (1980) suggest that teachers often focus on content chunks, rather than learning outcomes desired when preparing lessons. In addition, external mandates may affect teacher decisions, planning, and selection of materials. They may routinize teaching to deal with these demands (Cuban, 1986). However, other studies suggest that experienced teachers form expectancies of student achievement and that these predictions help drive decisions regarding selection of materials, pacing, depth of coverage, and level of mastery desired. This assists the teacher in managing the uncertainty of the classroom by categorizing, and sometimes grouping students by anticipated performance levels (Cooksey and Freebody, 1985).

While presenting lessons, teachers continue to make interactive decisions and to adjust instruction because of student responses they are receiving. Stiggins (1992; 1994) adds maintaining activity flow as well as surveying and accommodating for students' prior knowledge, to the list of interactive decisions made during teaching. Finally, teachers judge the instructional match of a particular text or strategy to students' needs as the lesson proceeds.

Early studies of teacher expectancy and prediction focused on the accuracy of their judgments. This began 40 years ago with the investigation of kindergarten and first grade teachers' accuracy in predicting first grade reading success (Henig, 1943; Kottmeyer, 1947). The predictive validity of these ratings equalled or exceeded standardized test scores. This prediction accuracy was later documented at the elementary level, (Doherty and Conolly, 1985; O'Connell, Dusek and Wheeler, 1974) and extended to long range...
predictions over several years (Ebbesen, 1968; Keogh and Smith, 1970; and Stevenson, 1986).

The best known of these early studies (Rosenthal and Jacobson, 1968) suggested a causal link between judgment and subsequent achievement. This is the first of the inquiries focusing upon teacher bias. They sought to investigate how a student's achievement is mediated by teacher decision making that was based on extraneous and/or noncognitive student cues. These included race, social class and gender. The results of gender bias studies are not uniform. Wiley, Eskilson (1978) and Hanes (1979) reported nonsignificant findings while Tobiessen, Duckworth, and Conrad (1971), Keogh and Smith (1970), and Stevenson (1976) reported gender bias in favor of girls. Time of the year or halo effects were reported by Guskey (1975) and Sullivan, Smith and Lopez (1989).

The ability of teachers to identify at-risk students correctly, and whether these judgments are mediated by bias, has been an active area of inquiry. Kagan (1988) reported that identification of at-risk students is a function of context cues such as the composition or variance of a given class and behavioral norms within that class. O'Connell, Dusek and Wheeler (1974) suggested that academic rather than social criteria inform decisions, while others support hypotheses that these decisions may be biased by ethnicity or socioeconomic status (Cooper, Baron and Lowe, 1975; Doherty and Connolly, 1985; Goodwin, 1969; Payne, 1989; and Rist, 1970). Ability and effort were reported as a basis for forming judgments (Cooper and Burger, 1980; Rosenthal and Jacobson, 1968) while labeled status was reported by several others. This included labeling for special education as well as a labeled status as a retained student. In both instances teachers reported differential expectations for achievement of labeled students who exhibited similar behaviors to nonlabeled students (Foster and Yesseldyke, 1976; Foster, Schmidt and Sabatino,
Noncognitive and sometimes extraneous variables affect teacher thought and behavior. This in turn impacts decisions regarding grouping, materials, strategies, pacing, etc. It appears these decisions may be mediated by cues while the teacher is unaware of their observed differential behavior (i.e., assisting students perceived to be at-risk in decoding unknown words in context while withholding such assistance for students perceived to be grade level and above). These students might receive extended wait time. The implicit expectation is that one group will not figure out the word and must be protected from frustration while others possess the tools and will elicit a correct response if given time.

While teacher prediction of achievement appears inextricably bound with teacher expectancy, the focus of the study presented here is upon the nature of data sources which inform teacher judgment. Clearly, data sources are not independent of classroom contexts in which they are gathered, documented, and valued. These contexts often yield up as much data about teacher cognition as they do about data sources.

Stiggins and Conklin (1992) lament the paucity of assessment studies conducted within the context of classrooms. Yet only teachers in this interactive environment may consider a longitudinal sample of a student's reading behaviors.

Our understanding of the nature and quality of classroom assessment environments and the task demands of classroom assessment are in their infancy. Further, we have very little understanding of how to prepare teachers to meet these demands (p. ix.).

**Novices and Experts**

As noted, the 1980's produced cognitive researchers such as
Resnick and Glaser who undertook to discover the processes involved in generating a competent performance. Previously, researchers had investigated the performance of experts. They discovered experts process incoming data in novel ways. Some information is "chunked" in order to process less but larger units of data in order to simplify processing. Miller (1956) and deGroot (1965) as reported by Bereiter and Scardamalia (1992) described experts as superior in their ability to recognize patterns and thus feed only plausible solutions to the brain. Hensley, Hayes and Simon (1977) and Robinson and Hayes (1978) as reported in Bransford and Vye (1989) explored the strategies experts use for solving problems. They found experts quickly categorized a problem and thus narrowed the range of strategies.

Berliner (1986) and Leinhardt and Greene (1986) studied expert teachers and found they could quickly infer what was going on in a classroom context from glancing at a picture of a classroom. Schneider and Shiffrin (1977) as reviewed by Bereiter and Scardamalia (1992) affirmed the automaticity of processing by experts. Novices, on the other hand, may feel overwhelmed by the amount or complexity of the data. Chi, Feltovich and Glaser (1981) are reported in this same review to note that this rapid pattern recognition is organized around principles within the expert's field. Bereiter and Scardamalia sum up the expert studies with this maxim: "novices think; experts know" (p. 522).

A related topic studied by cognitive theorists including Resnick (1987) concerned the acquisition of new learning. One passes through the stages of novice, experienced practitioner, skilled practitioner, and finally to expert. Expertise, they concluded, cannot be transmitted. Instead, new knowledge must be constructed. For teachers who are learning the art of teaching and decision making about pupils, this construction of schema or meaning may take time.
Gaines and Davis (1990) in investigating the accuracy of teachers predicting student achievement on standardized tests, did not find significant improvement in prediction accuracy as a function of experience until 14 years. This suggests that rapid pattern recognition may be a function of experience. Teachers must perceive patterns of learning or behavior repeatedly before they may "chunk" this information to be processed automatically.

Harp (1993) also legitimizes intuition as a valid way experts know who has learned. The caveat here is that intuition must be based on informed observation and knowledge of a child's learning. Finally, Johnston (1987) identified characteristics of expert evaluators. The first of these is recognizing patterns. While a novice is aware an error has been made, the expert categorizes the error (i.e., self-correction). Secondly, experts have procedural knowledge. This includes managerial skills which free the teacher's time from discipline to focus upon observation, recording and interacting. Also included here is the ability to schedule and capture desired data. Thirdly, the expert is a good listener. They understand and hear the child constructing meaning from text. Experts emphasize process over product. Finally, other criteria included are child advocacy, teacher ownership of assessment, and promotion of student self-assessment.

Teacher Change

Borko (in press) presents a paradigm for teacher change building on the cognitively framed staff development work of Richardson (1992). First of all, inservice must present new knowledge along with the call for revision. Teachers must have the new skills necessary to implement the change. Secondly, they must be allowed to confront their beliefs and attitudes in the light of the
call for change. Teachers confront change in a variety of ways. First, some filter change through their set of beliefs and practices. They adopt some new concepts and adapt others. These are the accommodators. Second, some persist in their own practices. This is called anchoring. Finally, some pick elements of a new paradigm to incorporate into their existing practices. These are the assimilators. Borko reported teachers do not acquire new knowledge as discrete skills. Rather they build on their understanding about learning and teaching in contextual units. Thus, attempts to present knowledge must be situated in classroom contexts in order that teachers may see its application and integration.

Another tenet of teacher change literature is that change takes a long time. Hence, a three month series of inservice presentations is considered a very short time for any meaningful change to occur (Borko et al., in press). Inservice on running records, presented over the course of an entire year, however, did cause teachers to report that they felt this new knowledge helped them gain information that told them more about their students.

Teachers are more likely to change or acquire new skills if they perceive teacher ownership in the process (Borko et al., in press). Externally mandated change is often resisted, ignored, or assimilated. Finally, teachers often lack experience in the area of assessment and therefore, gravitate toward instruction where they feel more confident and competent. Stiggins and Conklin (1992) attributed teachers' discomfort with a lack of basic assessment literacy needed to implement change. Borko et al. (1992) concluded that it is logical that change in teacher behavior in the area of assessment will be slow and occur over an "extended period of time."

Data Studies

The first major descriptive study of classroom assessment
contexts was completed by Jackson (1968). Before that time all major studies of classroom assessment focused upon the "hard data" of standardized tests. Jackson first recognized the maxim documented in expectancy prediction research (Brophy and Good, 1974): teachers can predict with greater certainty a student's level of success on outcome measures closest to the instructional task. Similarly, the data from performance and observation are preferred by teachers over externally mandated criterion referenced or norm referenced data. When making instructional decisions, Stiggins and Bridgeford (1985) and Salmon-Cox (1981) reported that teachers' highest frequency of data is observation. In Dorr-Breme and Herman's study of 475 elementary teachers, they reported that 95% of important decisions were based on "my own observation of a student's classwork" (p. 36). The caveat here is that this same research found that the great majority of teachers had had no assessment training. Further, teachers identified their more experienced colleagues as their source of assessment strategies and guidance. Yet, a further finding of the study was that teachers engaged in minimal collaboration in developing assessments. In fact, teachers are more likely to share behavioral data than classroom performance information (Pallas, Natriello, and Riehl, 1991). Thus, the expertise of these competent professionals may not be effectively utilized.

Documentation of data is another focus of these data studies (Stiggins and Bridgeford, 1985; Stiggins and Conklin, 1992). Forty percent of teachers surveyed in the first of these studies reported relying upon memory to document observation or performance assessments. In addition, the tenets of assessment such as informing students of criteria, grading procedures, etc., were rarely followed in either study.

Despite their reliance upon teachers' self-reported data, which may vary from their daily practices, surveys have persisted as the
predominant research design to investigate this area of study. In a survey study of 206 teachers, Barry (1992) reported only 50% of the teachers who identified themselves as whole language teachers accurately described their actual approach to reading. Thus, it appears that while teachers may report transition from a basal program (and may actually incorporate whole language strategies), they may continue to rely heavily on the basal as their primary mode of instruction (Pryor, 1992). Wood (1993) found an inverse relationship between a teacher's reported orientation toward whole language and reliance upon basals, phonics, or testing. However, these whole language teachers still conceded that grades on report cards were based on tests or seatwork. Portfolios were reported as important but were not used for grading purposes. Barry (1992), Pryor (1992), Stiggins and Conklin (1992), and Borko (1993) reported on the importance of word recognition and oral reading. However, they also pointed out that methodology may drive assessment practices. In addition, although teachers stated they valued oral reading, they failed to articulate either criteria for valuing or procedures for documenting (Pryor, 1992).

Pryor (1992) and Stiggins and Conklin (1992) reported that a teacher's belief system affected which data were gathered and valued. However, because of many external factors, the beliefs that teachers report and implement in their actual classroom behavior may vary. Thus, it appears that survey and interview data must be interpreted with confidence only when multiple sets of data are collected. Preferably, data should be triangulated with observation, document inspection, or think alouds. Indeed, Stiggins and Conklin (1992) recommended the rigors of ethnographic research.

Besides data sources, documentation strategies, and influence upon the teacher's development of assessment strategies, several other variables appear in the literature as influences upon the
assessment environments in the classroom. Social factors are explored by Airasian and Madaus (1983) and Salmon-Cox (1981). The latter reported the interaction of social variables, level of school, and assessment. Elementary teachers appear to place more importance upon social variables such as behavior, cooperation, and attention to task. Frairy, Cross and Weber (1992) reported teachers' belief systems affected the use of noncognitive data in grading. September "sizing-up" strategies that remain stable over time were reported by Pryor (1992), Stiggins and Conklin (1992), and Rist (1970).

Dole, Duffy, Roehler, and Pearson (1991) studied interactive decision making. Teachers were observed to rely upon students' oral responses and in fact, altered their behavior based upon this feedback. Incorrect responses led to "response elaboration" or alternative representation, while teachers responded to correct answers with decreased scaffolding. That is, teachers appeared to adjust the amount of structure they provided for students to construct meaning from text. This reduction was based upon their perception of a student's needs. Data for this interactive decision, besides the accuracy of responses, might be the teacher's judgment regarding student's prior knowledge on the topic, their word identification skills, oral language skills, and the student's prior success in comprehension.

This interactive model suggests that teaching reading and assessing reading, as well as reading itself, may be a constructivist activity. During the process of leading a discussion of a story, the teacher receives multiple pieces of data about a given student's learning. These might include oral responses, the quality of oral language, body language, behavior, oral reading errors, etc. The teacher then combines this information with knowledge of the student's past performance and needs, home background, and labeled

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
conditions such as retainee, Learning Disabled, or Attention Deficit. The teacher's response to the student, therefore, may be mediated by these data as well as exemplars for performance in this area.

**Conclusion**

From the review of the literature several questions emerge on key issues surrounding data gathering and valuing by teachers to make decisions in reading. The first set of questions is concerned with the nature of data collected.

1. What is the mix of formal to informal data collected?
2. What percentage of data is interactive (i.e., collected as students respond during teaching)?
3. What percentage of data is noncognitive data?
4. What documentation strategies exist to capture data?
5. What are the external mandates that teachers face for accountability, and how do they affect data collection?
6. What evidence is there of whole language assessment data as recommended by Harp (1993) and Goodman (1989) including performance assessments, checklists, anecdotal records, miscue analysis, etc.?

A second set of questions concerns the belief system of the teachers. They are presented below.

1. Where are these teachers on the change continuum between basals and whole language? How does this affect their data collection?
2. How did these teachers acquire their assessment expertise? What influenced them? How has their collection or valuing of data sources changed over the years?
A third set of questions examines the judgment of these teachers. They are presented below.

1. How do teacher's self-reported pattern of data collection compare with actual frequency of data sources tabulated?
2. How are September sizing-up decisions made? Which data inform these decisions?
3. How do teachers chunk information as they process data?
4. What is the role of noncognitive data in decision making about reading?

Finally, there are questions about the improvement of practice in this area. They include:

1. What recommendations do teachers offer to train new teachers to become reading assessors?
2. What recommendations do teachers offer regarding inservice issues in authentic assessment?
3. What concerns do these teachers express about the transition to alternative methods of assessing reading?
CHAPTER 3
PROCEDURES

Introduction

This chapter includes a discussion of the research method and plan. The plan includes an in-depth description of the setting for the study, a description of how subjects were selected, and how entry was secured. The chapter concludes with details of data collection procedures and data analysis.

Research Method

The questions for this study focus upon data sources and reflection by teachers about the data sources which inform their decision making in reading. These questions drive the study into the vertex formed by behavior intersecting with context. Previous studies have demonstrated that teachers know a great deal about their students. This inquiry asks what data assist them in coming to that state of knowing and how they developed strategies for collecting and valuing information about students in order to know them.

Seidman (1991) suggests that many questions in education fall into the social science arena rather than that of natural science. According to Seidman, these are questions that center on reflection and a desire to place meaning with behavior. His assumption and that of many qualitative researchers is that human beings are able to reflect upon their behavior and give it meaning that cannot be observed or intuited by another. In-depth interviewing allows this study to gain a window into teacher thinking about their purposeful behavior in a classroom context. The quantity of data produced by multiple interviews allows some confidence in their self-reported data sources. As important, repetitive interviews allow the
researcher to compare teachers' reported decision making strategies with those distilled from their profile. Finally, a rich, verbal description of teacher thinking about assessment and data sources that yield information about students' learning, provides an important resource for those who plan and implement teacher training and preservice education. The fact that these data are derived from a context that is in a state of transition makes the data more valuable to those who seek to implement meaningful reform in reading. It is important to understand the needs of experienced teachers as they find their way through the paradigm shift from a deductive to inductive method of teaching reading to young children. The case study methodology using in-depth interviewing is appropriate, therefore, to study these questions about what teachers do, know, and think (Patton, 1990). This information is triangulated in the study through the use of think alouds, classroom observations, and inspections of documents.

Research Plan

Setting and Participants

Participants in the study were seven second grade teachers in the York County Public School System in Virginia. They were chosen from those second grade teachers in the system who had ten or more years of experience in the primary grades and the recommendation of their principal and reading specialist. Principals and specialists were asked to nominate teachers who demonstrated exemplary ability to use all types of assessment data to describe student achievement accurately. Participating teachers were from five elementary schools which represented a wide range of socioeconomic levels and student achievement.

The Schools

The setting of the study was the County of York, a part of the
Jamestown-Williamsburg-Yorktown historic triangle on the Virginia peninsula. This 45 mile long county varies from rural to military to suburban areas bordering on NASA. There are 10,000 students enrolled in the public schools in the county and ten elementary schools. Of the five schools involved in the study, three were schools where Chapter I programs were mentioned. Determination of eligibility for this remedial program is based upon the total percentage of free and reduced lunch students in that school.

The language arts instructional program in the school district consisted of an adopted basal text, Harcourt Brace. Use of accompanying practice materials, including workbooks, varied from school to school. The criterion referenced tests that accompanied the basal were mandated by the district to be given a minimum of twice a year. However, beginning with the school year 1990-91, a countywide Developmental Primary Committee began studying appropriate developmental practices. The impetus of this work and an ongoing program of inservice on literature based classrooms began a paradigm shift away from sole reliance on the basal and use of workbooks. Novels, trade books, big books, etc. became either a supplemental or optional instructional material. However, use of the basal test continued.

Writing instruction followed the stages of the writing process. At all grade levels, students participated in planning, developing rough drafts, editing, revising, and preparing publishable products. Language experience, journal writing, grading by domains or traits, peer revision strategies, invented spelling, and reading responses were common elements of the writing program. Teachers were encouraged but not required to keep writing folders with representative samples of student writing, including the steps of the process for a given piece of writing. Writing folders were maintained for each student in the majority of classrooms until 1990.
There are several reasons why this school division appeared to have implemented reforms in the writing instructional program ahead of the reading program. First of all, this division had participated in the Eastern Virginia Writing Project throughout the last ten years. Those teachers who attended the summer institute were thoroughly trained in aspects of the writing process.

Another influence upon reform was the Virginia Literacy Passport Test in Writing. This performance assessment is domain scored and is given as a predictor in the fourth grade. The gatekeeping administration of this test follows in the sixth grade. That is, students who fail to demonstrate competency on the test receive remedial instruction in seventh and eighth grades and are retested. Students must pass this test to be classified as ninth graders. The accompanying inservice for teachers since its inception in 1987 has included planning strategies for writing including mapping, graphic organizers, brainstorming, etc. In addition, teachers have received instruction in writing domains, writing across the curriculum, and reading response techniques. They have received training on criteria for domain scoring of writing through the use of anchor papers.

This study took place at the end of a fifteen year period during which minimum promotion standards had been in effect in reading and math at all elementary grade levels. These standards in reading were tied to the criterion referenced tests that accompanied the adopted basal.

Grouping for instruction was accomplished through the basal reading levels with homerooms typically reflecting two or three reading groups. The exceptions to this were Elaine and Roberta, participants who taught one period of homogeneously grouped students in 1990-91 in a team-teaching setting.
Subjects

The seven teachers who participated in the study had from ten to 35 years experience. They were white females who ranged in age from 38 to 65. All but one had taught in their respective schools for four or more years. The exception was moved to another school for the fall of 1991, and requested a transfer to yet another school in the fall of 1992. Thus, interviews with this teacher were conducted at two different schools.

Those teachers with an undergraduate degree only, held this degree in elementary education. The subject with the least experience had entered the work force after raising her children. She attributed her knowledge of developmental levels to watching her own children grow and mature. She stated that she continued reading journals on child development and early childhood education during this period. This was the only reference to professional literature during the study. This subject did substitute teaching for four years before returning to teaching. She stated this gave her an opportunity to adopt strategies and techniques that she observed worked well in classroom context(s). Finally, this subject made the only reference in the study to school-based staff development. She stated that inservice on assessment contributed to her knowledge and expertise in this area.

It should also be noted that the subject described above participated in a previous study on teacher prediction accuracy (Gaines and Davis, 1990). Her ability to predict student achievement on standardized tests in reading exceeded all other second grade teachers in the school division.

Another subject had taught primary grades only for over 20 years in the same school. She described herself as having a "hands-
on" classroom. She was comfortable with movement of children, cooperative learning, creative projects, and stated a strong preference for literature rather than basal materials. This subject described her own love of reading and books. The physical arrangement of her room included pillows, stuffed animals, and games. Books the children had written were on display in the room. Children were observed to engage in recreational reading on the floor with friends or with a pillow alone in a corner.

Another subject had taught 35 years in many settings and grades as her husband's work required frequent moves. In addition, she had spent one year teaching in an in-school suspension program. Suspended students did not arrive in her classroom with any assignments; therefore, she had to assess skills and needs quickly and provide materials and activities. She attributed her powers of observing and diagnosing to this experience. Her classroom observed in this study was unique because of additional adults observed there. Field notes indicate these people included an instructional aide assigned to a special education student, and an elderly volunteer from a nearby church.

One subject had taught kindergarten for ten years before switching to primary grades. She had worked with three other kindergarten teachers and described peer teacher mentorship as a strong influence on her development. This teacher taught in a school with a large number of free and reduced lunch students.

Yet another subject had only taught primary grades. She characterized herself as a structured teacher and stated a preference for basal materials. Her students were primarily military dependents.

One subject had a Master's Degree in special education. She had taught self-contained classes with learning disabled and emotionally disturbed students. She had also conducted educational assessments for eligibilities which included standardized testing and
curriculum based assessments.

The final subject in this study had taught upper grades as well as primary grades. She was enrolled in a graduate program in reading during this study. This teacher provided the most elaborated responses in the study. She felt comfortable adding information to her responses, or commenting on related topics.

Gaining Entry

The dissertation prospectus was submitted to the Human Subjects Review Board and was granted approval. A Request to Conduct Research was submitted to the York County Director of Program Evaluation. This request was subjected to a three member blind review and was subsequently granted approval. Due to a change in administrative organization in the school division, the Director of Elementary Schools was assigned to direct external research. Therefore, a letter describing the proposed research and referencing the prior approval was submitted to this director. Approval was granted. This director forwarded her own letter to the division elementary principals describing the project and granting access at the school level, contingent on principal approval/recommendation. The researcher then contacted each elementary principal. There were eight second grade teachers in the division who met the criteria for experience. Seven were recommended for participation in the project.

The researcher was personally acquainted with six of the seven subjects. However, her duties as administrator and researcher did not overlap. For example, while she evaluated teachers, she did not evaluate any of the subjects.

A proposal was also submitted to the Research Committee of the Virginia Educational Research Association with a request for grant funding for the purpose of compensating informants in the study. The proposal received a grant award of $500.

The researcher then contacted each subject, informed them of
the nature of the project, the extent and duration of the interviews, observations and videotaping, their rights and protection, and the nature of compensation. A letter describing these features of the project was then forwarded to each participant. Subjects were free to decline to participate. Teachers were assured that pseudonyms would be used in written reports and that tapes would be viewed and transcribed only by the researcher (See Appendix A). Initial teacher profiles completed at the conclusion of the interviews were provided to each participant with a request for feedback. One subject received a complete transcription of an interview at her request. She then provided further elaboration of responses and information about the context of certain statements.

Data Collection

The researcher employed triangulation of the data in this study for several reasons. First, multiple sources provided added confidence in the validity of trends in the results. Secondly, the classroom context is highly complex. Data sources appear to overlap and interact. Therefore, it is important to gather information in a variety of contexts. However, an important caveat here is that the ultimate goal was not multiplicity of data (Seidman, 1991). Instead, knowing was investigated from a variety of perspectives. These included observing, listening, inspecting, probing, and comparing. Thus, triangulation was achieved through use of interviews, document inspection, classroom observations, videotaping, and teacher think alouds.

Interviews. Interviews were conducted during the 1993 calendar year. Six semi-structured interviews per teacher focused upon how they knew who was learning in their classrooms, how needs groups were formed, how grading was accomplished, and what type of accountability was necessary to communicate effectively with parents about their child's learning. Teachers were also asked to describe their own
development as assessors and "kid watchers," and to make recommendations regarding assistance programs for beginning teachers.

Interviews were conducted by the researcher at the participant's school, after the regular school day. The interviews were semi-structured using an interview guide (See Appendix A). Probing questions or requests for additional information were either accomplished through repetition of the informants response or framed through an analysis of their previous responses. Data sources were deliberately not specified in the interviews unless mentioned by the informant to avoid leading their responses. Thus, questions such as "do you value student's attention to task?" were not included. However, if a respondent had talked at length about the interaction of attention to task and self-monitoring of comprehension by a student, for example, this segment might be read back to the informant with a request to provide additional information on this topic.

Observations. One classroom observation was conducted for each subject. Subjects were asked to select a classroom segment that offered them a good opportunity to gain information about their students' literacy learning. Although all subjects requested additional information about the setting to be observed, or asked that a given setting be specified, only the above statement was reiterated. The purpose of providing an open-ended global request for an observation setting was to gain information about teacher thinking about which settings provide them with assessment data. Observations were approximately 45 minutes in length. They were all accomplished during the spring of 1993. Field notes were recorded by the researcher during the observations and afterward, as well as scripting of teacher behavior.

All interviews were audiotaped. In addition, the researcher
made notes during the interview and post-observation notes after the interview. The latter recorded the researcher's thoughts, identified themes that appeared to emerge for a given respondent, as well as topics that needed to be reintroduced in subsequent interviews. For example, Amy was the only subject who mentioned checklists and performance assessments as one of her assessment strategies. Therefore, a note was made of this in order that documentation could be inspected and this topic could be reintroduced if appropriate. This topic was reintroduced during the interview on grading practices.

**Think Alouds.** The researcher explained the process of think alouds to each participant and asked them to have a ten minute videotape made of a classroom instructional segment in reading. They understood that the format of the succeeding interview would be to watch the videotape with the researcher. Periodically, the researcher or the teacher would stop the video in order to allow the teacher to talk about what she was thinking, feeling, judging, etc., as the lesson progressed. Think alouds were also audiotaped, transcribed, and analyzed as interviews. The researcher made post-observation notes comparing observed strategies to self-reported ones made during interviews.

**Documents.** One interview focused upon samples of student work. Teachers were asked to bring several samples of work for more than one student. They then reflected orally on what information they gained from work samples. The quantity and types of work samples provided by the informants varied widely. In addition, teachers often mentioned documents in the course of interviews and these were requested. For example, Amy talked about a type of checklist she had created to record an informal performance assessment of different skills.

Other documents which were inspected included a sample (usually...
two or three per informant) of cumulative folders, reading folders (including criterion referenced basal tests), and the teachers' grade books.

Teachers' reflections about documents were audiotaped in the course of these interviews. Transcriptions were analyzed as interviews. However, hard copies of work samples were retained for comparison and reflection.

**Data Analysis**

**Rationale**

Patton (1992) makes the point that there is no clear demarcation between data collection and analysis. In the process of the former, trends in the data appear and ideas about analysis emerge. Therefore, initial focused questions were soon augmented by a dichotomous division in the data between cognitive and noncognitive data. Initial attempts, however, to analyze the data with these trends in mind demonstrated clearly that cognitive and noncognitive data were not valid divisions. Scotty talked about the child who could not monitor his own comprehension during testing sessions because of his distractibility. Therefore, she provided the modification of testing this child individually, and asked him to read aloud to himself. In this way, he cued himself to monitor for meaning. She reported marked gains in comprehension scores. This information is neither cognitive nor noncognitive. Instead it demonstrates clearly how behavior (distractibility) can be given meaning (distractibility depresses the student's ability to monitor his own comprehension), that has important instructional implications (cuing strategies improve the ability to monitor comprehension), in a particular context (testing) by an experienced teacher. This is clearly different from the teacher who includes effort in her grading scale. In this latter case the teacher has introduced a noncognitive
factor into a measurement of learning outcomes; therefore, it becomes a source of error.

What emerged in the enormous volume of data in 56 files was repetition of data sources grouped under specific headings of oral reading, oral language, oral responses, behavior, work samples, etc. In addition, teachers' comments about their individual decision making styles began to emerge. Indeed, Patton admonishes the qualitative researcher to "observe their own processes even as they are doing the analysis" (1990, p. 372). Initial analysis, therefore, took the form of frequency counts of data sources for each informant across all sources, and frequency counts of data sources across all informants. Individual profiles were completed and followed by cross-profile analyses. To provide member checks, individual profiles were presented to each teacher with a request for clarification and feedback.

Prior to the study, the researcher anticipated that categories of responses might follow the kind of data recorded in an informal reading inventory or miscue analysis (Goodman, Watson and Burke, 1987; Johnson and Kress, 1965; Farr and Carey, 1986). These would include word recognition in isolation and context, features of oral reading, analysis of miscues, and oral and silent comprehension. In addition, from an awareness of developmentally appropriate practices in the literature (Harp, 1993; Adams, 1990; Trail, 1993; Clay, 1982; Strickland, 1989; Glazer and Searfoss, 1988; Hill and Ruptic, 1994; Adams, 1990) as well as that of Strickland and Morrow (1989) and Routman (1991), the researcher became aware of data typically recorded in a language-based classroom. These included retellings, responses to literature, journals, analyses of invented spelling, reading logs, etc. The researcher's experience in monitoring student progress at the school level on criterion referenced basal tests, made her aware of skills tested on these instruments.
Finally, several previous data studies provided examples of both data and decision making categories. Dorr-Bremme and Herman's study (1986) included survey data with categories chosen at the beginning of the study. Teacher opinion was included in this study as a source of data. The combined categories of teacher observations/opinions received the highest reliance rating for all types of decisions. Thus, teachers recognized and self-reported that their intuitions and "gut level" feelings strongly affected their planning, grouping, grading, and modifications for students. Stiggins and Conklin (1992) provided extensive lists of categories and decisions in their participant observer instrument, designed to consider all aspects of the assessment environment in the classroom.

Other categories emerged as the data analysis proceeded. Categories such as oral language emerged from the data when it became clear that subjects differentiated between the correctness of an oral response and the quality of a student's oral language. Finally, strategies for valuing data emerged such as comparison, triangulation, observation, pattern recognition, and an awareness of deviation from trends.

Qualifications of researcher

The researcher was qualified to conduct this research due to her graduate degree in reading and 20 years of experience as a reading specialist. In this capacity the researcher had ongoing dialogue with teachers about their students, their progress, the data they collected, and externally mandated data. The researcher tested students for placement and for diagnostic purposes and served on child study, eligibility committees, and chaired student assistance teams. Teachers were assisted in interpreting the results of standardized test data at the school, classroom, and pupil level for purposes of program evaluation and instructional monitoring and
revision. The researcher completed a 200 hour internship in the Program Evaluation Department of the York County School Division. In addition, the researcher had completed a previous quantitative study (Gaines and Davis, 1990) on teacher prediction. Follow up interviews conducted provided this researcher with experience in framing interview questions as well as a sensitivity to issues regarding interview research.

In summary, the researcher's professional background and experience, coupled with an awareness of the literature, were aids in coding text information and recognizing additional categories of data as they emerged in analysis.

Procedures

All transcriptions were entered into a qualitative data base (Padilla, 1991). Tagged codes were assigned in an open coding system to units of text that specified data sources or stated judgments. Inspection was then made for emerging trends in the data and text information was examined for recurring data themes (i.e., word recognition, behavior and work habits, oral language, etc.). These were tallied for each participant and reported in percentages of total data sources. A mean percentage was reported for each category of data.

A profile of assessment strategies was developed for each teacher by combining their self-reports, observations, and inspection of documents. In addition, comparisons were made across profiles for common trends which appeared to emerge in sources and valuing of data.
CHAPTER 4
RESULTS AND DISCUSSION

Introduction

This chapter includes the results of data analysis on teacher interview data, observations, and document inspection. A summary of data responses for all subjects is presented. This includes only the highest frequency data categories. Individual subject profiles are then presented which summarize the frequency of each type of data source mentioned in all sources. Quotations are drawn from the interviews to illustrate teachers' strategies in collecting these types of data, as well as valuing the data and utilizing them for decision making. Next, common themes of data are discussed including observational data, comprehension, oral language data, prior educational data, work samples, tests, grades, and creative writing. Themes of data vary among profiles; therefore, inclusion of a discussion regarding a theme is a function of the frequency of that category of data within an individual profile. Class makeup and methodology are addressed as variables relating to the data profile. Unique influences upon the data profile for each teacher are covered. Finally, the teacher's self-assessment of her growth as an assessor and her awareness of her own style is presented. This is accompanied by any articulated concerns and recommendations for teacher training in this area.

It should be noted again that all quotations of subjects were recorded on tape and transcribed verbatim. The use of ellipsis marks [. . .] does not indicate that material has been deleted. Rather, it denotes that subjects' statements were most often in partial thoughts expressed in incomplete sentences.
Cross-case analyses are presented and discussed in the context of emerging themes. The findings of the study are then considered in light of preceding theory. Finally, a typical profile is presented at the conclusion of the chapter that integrates the common data sources and strategies of the seven subjects. This constructed theory is compared with current cognitive constructivist views of literacy and recommended best practices in assessment. Discrepancies are noted and discussed.

**Summary of Data Sources**

A summary of data percentages for the highest frequency categories is presented in Table 1 for all subjects. It is introduced here in order to provide a point of comparison for individual profiles and the cross-case analyses which follow. Interactive data and other aggregates of data usage are not presented here as they emerged during data analysis and are discussed at the end of this chapter.
**Table 1**

**Summary of Percentages for each Data Source**

<table>
<thead>
<tr>
<th>Source</th>
<th>Roberta</th>
<th>Amy</th>
<th>Elaine</th>
<th>Scotty</th>
<th>Stacy</th>
<th>Kitty</th>
<th>Betty</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>277</td>
<td>261</td>
<td>260</td>
<td>329</td>
<td>275</td>
<td>165</td>
<td>193</td>
<td>251</td>
<td></td>
</tr>
<tr>
<td>W.R.</td>
<td>15%</td>
<td>12%</td>
<td>11%</td>
<td>11%</td>
<td>18%</td>
<td>21%</td>
<td>8%</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>C/D</td>
<td>7%</td>
<td>9%</td>
<td>8%</td>
<td>10%</td>
<td>8%</td>
<td>7%</td>
<td>12%</td>
<td>9%</td>
<td>2%</td>
</tr>
<tr>
<td>Obser</td>
<td>10%</td>
<td>4%</td>
<td>2%</td>
<td>5%</td>
<td>9%</td>
<td>6%</td>
<td>13%</td>
<td>7%</td>
<td>4%</td>
</tr>
<tr>
<td>Behav</td>
<td>6%</td>
<td>7%</td>
<td>8%</td>
<td>11%</td>
<td>4%</td>
<td>5%</td>
<td>8%</td>
<td>7%</td>
<td>2%</td>
</tr>
<tr>
<td>Oral Resp</td>
<td>9%</td>
<td>7%</td>
<td>7%</td>
<td>5%</td>
<td>5%</td>
<td>4%</td>
<td>5%</td>
<td>6%</td>
<td>2%</td>
</tr>
<tr>
<td>Test</td>
<td>10%</td>
<td>3%</td>
<td>8%</td>
<td>6%</td>
<td>4%</td>
<td>3%</td>
<td>7%</td>
<td>6%</td>
<td>3%</td>
</tr>
<tr>
<td>Oral Lang</td>
<td>8%</td>
<td>9%</td>
<td>4%</td>
<td>8%</td>
<td>3%</td>
<td>5%</td>
<td>3%</td>
<td>6%</td>
<td>2%</td>
</tr>
<tr>
<td>Comp</td>
<td>1%</td>
<td>5%</td>
<td>8%</td>
<td>9%</td>
<td>8%</td>
<td>6%</td>
<td>3%</td>
<td>6%</td>
<td>3%</td>
</tr>
<tr>
<td>Work</td>
<td>4%</td>
<td>3%</td>
<td>7%</td>
<td>3%</td>
<td>1%</td>
<td>7%</td>
<td>3%</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>Writ.Lang.</td>
<td>3%</td>
<td>3%</td>
<td>6%</td>
<td>4%</td>
<td>2%</td>
<td>4%</td>
<td>0%</td>
<td>3%</td>
<td>2%</td>
</tr>
</tbody>
</table>

*Note.* Sources of data include: Word Recognition, Comparison of Data, Observation, Behavior, Oral Responses, Test Data, Oral Language, Comprehension, Work Habits and Written Language.
Roberta's Assessment Profile

Introduction. Roberta teaches a homogeneous group of at-risk students in a school characterized by many free and reduced lunch students. Her instructional style is oral and interactive. No paper-pencil activities were observed. She appears to value hands-on activities and student involvement. Her statements reveal that she believes reading must be made meaningful for students in order to motivate them. Therefore, one observes puppets, cereal boxes, and many everyday signs in her classroom. In addition, she has invited other adults to participate and assist.

Table 4.2 records the frequency of data sources stated in all interviews. The total number of data sources identified was 277. A percentage of total responses is reported for each data category.
Table 2

Data Frequencies for Roberta and Percentages of Responses by Category

<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Recognition</td>
<td>41</td>
<td>15%</td>
</tr>
<tr>
<td>Observation</td>
<td>29</td>
<td>10%</td>
</tr>
<tr>
<td>Basal Test Data</td>
<td>27</td>
<td>10%</td>
</tr>
<tr>
<td>Oral Responses</td>
<td>24</td>
<td>9%</td>
</tr>
<tr>
<td>Oral Language</td>
<td>21</td>
<td>8%</td>
</tr>
<tr>
<td>Comparison of Data</td>
<td>18</td>
<td>7%</td>
</tr>
<tr>
<td>Home Background</td>
<td>13</td>
<td>5%</td>
</tr>
<tr>
<td>Work Samples</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Motivation</td>
<td>9</td>
<td>3%</td>
</tr>
<tr>
<td>Interest</td>
<td>8</td>
<td>3%</td>
</tr>
<tr>
<td>Written Language</td>
<td>8</td>
<td>3%</td>
</tr>
<tr>
<td>Standardized Tests</td>
<td>6</td>
<td>2%</td>
</tr>
<tr>
<td>Ability</td>
<td>5</td>
<td>2%</td>
</tr>
<tr>
<td>Memory</td>
<td>4</td>
<td>1%</td>
</tr>
<tr>
<td>Prior Educational History</td>
<td>4</td>
<td>1%</td>
</tr>
<tr>
<td>Grade Level as Reference Point</td>
<td>4</td>
<td>1%</td>
</tr>
<tr>
<td>Comprehension</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Health Issues</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Peer Coaching Data</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Skills</td>
<td>2</td>
<td>1%</td>
</tr>
</tbody>
</table>

*table continues*
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data from Reading Specialist</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>L.D. Label</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>ADHD Label</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td>Anxiety</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td>Regression</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td><strong>Total number of data sources</strong></td>
<td><strong>277</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Note.*  * = less than 1%
Discussion of Roberta's Data Sources and Assessment Strategies

**Highest frequency data and class makeup.** The text from all of Roberta's interviews contained 277 references to data sources. Of these, almost 15% (14 statements) were related to word recognition. This may or may not be due to the makeup of her classroom. During the language arts instructional period at this school, teachers instruct a homogeneous group of students. Roberta's class is comprised of remedial, below grade level students with severe decoding difficulties. In fact, she characterized several of her students as not being able to read at all. However, it should be noted that word recognition concerns appear in all subjects' profiles while the class makeup of some subjects' classes is above grade level. Roberta describes the class' reluctance to read as a result of their history of failure. Therefore, the category of "willingness to read" may also be regarded as related to class makeup. Finally, she makes 13 references to home background. These are usually comments related to whether students receive reinforcement at home or whether they have been exposed to many or few experiences. These may be considered a function of her students' socioeconomic status which was primarily free and reduced lunch students.

Comparing Roberta's profile with Elaine's offers an opportunity to consider class makeup within a given school. Roberta and Elaine both taught at the school with the largest number of free and reduced lunch students in the school division. However, Elaine had a reading class of gifted and above grade level students. In comparison to Roberta, she made only five statements related to home background. Another significant difference between the two teachers is their reported reliance on observational data. Roberta made 29 statements (21%), while Elaine made only 4 (2%). In addition, Elaine stated her reliance upon work samples and stressed the importance of above grade
level students being able to produce written products independently.

**Comprehension.** Comprehension was mentioned specifically three times during all interviews with Roberta. Features included recall, understanding of word meanings, and cause and effect.

**Oral language and oral responses.** Oral language and oral responses were mentioned 21 and 24 times respectively and are consistent with the teacher's view of her teaching style as oral vs. written. "We read and reread... I have few work samples." This respondent appears to use oral language in order to assess what words may or may not be present in a student's receptive or expressive vocabulary. She then makes instructional revisions based on her perception of student needs in this area. For example, she used the word *supposed* and gave the sentence from text, "A dragon is supposed to be a monster." Her assessment revealed that this word was not in students' expressive vocabulary because they stumbled over repeating it after her. Further oral activities revealed that the word lacked receptive meaning for students as well. In reading *The Three Bears*, she noted that the word *cottage* lacked meaning for students. Her generalization from this discussion is that her students are unable to supply synonyms for many words encountered in text. She indicates she will develop oral language and creative dramatic activities to develop these. Thus, one is able to observe that this teacher has made the full circle from planning to implementation, to interactive observation and informal assessment, to decisions about instructional revision, and finally back to planning.

**Decision making style.** Perhaps the first thing that emerges from Roberta's interviews is the fact that this teacher is reminiscent. Thus, her responses are illustrated by stories of many children. It is as though she states a theory (i.e., a given set of data may be interpreted in a variety of ways), and then culls through her varied experiences to test that theory. She shares a judgment
she made that turned out to be based upon invalid data. She expands on an incorrect decision made regarding her own child. Finally, she restates her distilled wisdom in almost poetic terms:

. . .but I think about how they observed him. . .and how easy it is to make a mistake. . .when we don't know the real will of the child. I have to have data to tell people, but I do not use that as the answer to all.

This is the first incidence of a common theme of decision making. When she states, "I have to have data to tell people," she is most likely referring to externally mandated collection of data for accountability purposes. These data might include a criterion referenced basal test, for example. When she balances this with the statement that "I don't use that as the answer to all," one might infer she is referring to the second level of assessment in her classroom. That assessment program is informal, interactive, and relies mainly on observation, interaction, and intuition. It is rarely documented; hence, she has to have "data to tell people."

Consistency between self-reports and data profiles. There appears to be a consistency between her verbal reports of the types of data she believes she favors for instructional decisions, and the tallies of actual responses across interviews. For example, she states that she has fewer work samples than some other teachers because the children spend their time "reading and rereading." This is supported by the high number of word recognition responses (41), and the fewest number of work samples provided (two). She appears to use observation and analysis of response at every level of data with the possible exception of word recognition. For example, when students are working she moves about the room and observes who is not actually reading the material, who is off task, and who is copying from another student. This increases the reliability of decisions
she is able to make with work samples. When students are reading their stories to her she notes the text, whether they read what is written on the paper, affective responses such as whether the student is proud of his story as evidenced by him holding it up and saying, "I have more to read to you," and the approximations of spelling.

**Word recognition data.** When students read orally she reports focusing on "how smoothly" they read. She does not report attending to the kinds of miscues students make. In fact, her purpose for observing during oral reading is to discern who is not actually reading. Roberta discusses using choral reading and group reading as a strategy:

...So today I stood right by him as we were reading "church style" and discovered too that a lot of times...
...they'll hum... They will not say the words and I have to encourage them to move their lips and make the sounds of the words...

She states she is more concerned with children acquiring the global concepts of the rhythm and flow of language in order to use context clues than in focusing upon individual miscues:

...At the same time I want them to get the whole picture of the whole sound, the whole rhythm of a sentence rather than the spasmodic reading that they do...

**Written language.** In dealing with oral reading of creative writing, however, she examines both the approximation of the invented spelling and the ability to decode the word from the invented spelling to obtain information about the word attack strategies the student demonstrates:

...He wrote... all of them had to write the best way
they could. . . I didn't help them much with words. . . and
if you notice there was one word he couldn't read himself
after he had written it down. . . and it was the word
spring. . . about the fact that it was spring. . . and
the swan realized that he was a duck. . .

Her ability to describe the information she obtains from
miscues when reading students' creative writing and from invented
spelling, is evidence that she understands oral reading miscues from
text. Therefore, it is important to consider whether this teacher
processes that type of information automatically and therefore,
doesn't report it in an interview.

Observational data. This teacher actively collects data during
testing. She monitors for students who may be guessing, students who
may not be actually reading the test, students who are cheating, and
students who request to have words supplied. She states:

. . . When they take a reading test. . . this was so
clear. . . we discovered that some of these children
could not read. Now either they have been copying off of
other people's tests or they had forgotten what they had
learned. I think lots of times. . . it's "oh, I'll fill
this in. . . we'll fill this in. . . we'll guess at this
one." They don't really read. They're lazy. . .

Test data. Invalid test scores are a recurring theme for this
teacher. Her concern extends to both the tests she has administered
and the test scores that come from the previous teacher. She thinks
aloud about the hazards of grouping children for instruction with
incorrect data:

Well, I had him. . . he was misplaced. . . I didn't have
him at the beginning of the school year. . .how he
functioned in another group I don't know. . .but some of
these children. . . again. . . I think because they have
cheated so much. . . or copied. . . or compensated. . .
or whatever the word is. . . they have come out looking
good on tests. . . and they don't know how to read. . .
and the teachers discovered that they couldn't read. . .
and one by one, I'm getting children added to my list. . .

Finally, she puts her consideration of test scores in
perspective. It is evident from her comments below that she
understands the dual roles of assessment. She knows the data
collected in her classroom may be used not only for instructional
revision but also for accountability and program evaluation.

Now test scores. . . grades. . . are very important. . .
but they're not the whole story of a child. . . but we do
have to have some sort of measurement. . . to cover
ourselves legally. . . to say yes, we have covered this
amount of material; this is what we've been teaching. . .
but it's not the true value of what a child is doing,
but it is a measuring tool. . .

Prior educational data. There is inconsistency among her
comments regarding the use of prior educational data and the actual
frequencies of such data. She declares such data to be important in
decision making. However, only four statements support this.
Further, she declares that she is highly attentive to family
statistics. However, in another interview, she admits that she may
never read the entire cumulative folder (including prior report
cards). She apparently views data from the previous teacher in
another light as she reports:
...well, I do need to know... two of our teachers that had first grade are not here this year and I can't talk to them. I was thinking about ____, and I thought, "gee, I would like to talk to somebody"... you do go back and talk to different teachers who had children...

In another interview she makes a statement that will become a common theme in this study. She states that she does not consult prior educational data until she has made up her own mind about the child. This appears to concern her fear of expectancy bias. Her "sizing-up" strategies for September are unique. Sizing-up strategies refer to those strategies used by teachers in September to assess a new class, develop attributions about ability and achievement, and group for instruction. Her September strategies focus upon surveying the children's interests and what they want and expect to learn that year.

Grades. Grading is an important issue for this teacher. She recalls being instructed by a former principal to have many grades for documentation. Therefore, she reports grading everything including pretesting and introductory lessons on a skill. She groups grades by reading skills and prioritizes tested skills. She includes grades from basal tests. In describing her gradebook, she states:

...Well, usually, my grades were from when we began with a skill... maybe not with any teaching... but let's see how well... maybe a pretest, you know... I and then we can see how it would improve as time goes on. You just can't say, "Well, your child is a B student."
They would say, "Well, why is my child a B student? Why aren't they an A student? They read everything at home?"
Yes, I wanted a lot of grades... you've got to have
something to back what you are saying.

In summary it appears that Roberta's decision making style involves collecting informal data that may not be documented to inform her attributions for individual students. However, she does collect many grades and test scores due to either external mandates, or her own need for documentation and accountability. She appears to use triangulation of data to increase her accuracy. For example, she observes during testing and then compares test performance with daily grades and her observations and recall of performance in the instructional group.

Data collection and methodology. Several factors appear to affect the assessment context in this classroom. As stated, Roberta characterizes her students as remedial and reluctant readers with limited experiences and oral language. She describes children who "look at the ceiling instead of at the book," and who "do not make the connection between the words in the book and the words in their real world," and who are "too afraid of being wrong to respond." This reluctance to read coupled with behavior and distractibility concerns appears to drive instructional methodology choices for this teacher. Although she states that she favors whole language and multi text-materials, she uses the basal. She rationalizes the use of the basal to provide structure:

... This is what you do today... and this is what we do tomorrow... this is what page you're on... it gives definite structure... but if you do something a little different... they sometimes will fall apart...

Even her selection of choral and unison reading is driven by her need to provide focus. Focusing students' attention appears to be a response to behavioral concerns as she states:
because if you wait for your turn, you can, again, do a lot of different things. 'cause there are lots of things you can do with a pencil. there should be a book written on how many ways you can use a pencil.

Thus, the methodological decisions she has made because of perceived student characteristics and needs significantly affect the type of data she is able to gather on student achievement.

Growth of assessment expertise. This teacher is unique in describing her growth over her career in assessing student learning. Although she is insistent that expertise in assessment comes with experience, she describes an atypical beginning teaching experience where she did not have the managerial or behavior concerns that characterize most entry level teachers (Hollingsworth, 1989). "I just never had them [management problems]. I went into a classroom and I taught." This is interesting in that most of the literature reviewed on beginning teachers suggests that teachers cannot attend initially to the simultaneity of the classroom and therefore, they direct their attention to management. This teacher stated that even with no significant management problems, she did not regard herself as a capable assessor in reading until she had more experience in watching readers grow and learn. She responds as most other subjects in crediting peers who mentored her through her first few years of teaching as she reminisces below:

... and so they were telling me, "Now, this is so and so's child. ... I taught this father and I taught the child's mother". ... and they would give me a thumbnail description of these children and how well they learned. ...and I had to observe them. ...and every child was on a different level. ...but it wasn't like there were little groups. ... and little circles and things. ... up and
down the aisles they went... each child... O.K. ... helping this child... they helped this child... over here... over here... I'll never forget... their management... they still... those two teachers are such a blessing in my life as I think of them...

Written products. Papers provided included a child's spelling paper which Roberta analyzed to determine approximation of the word given. A workbook page was included. This teacher noted reversals, letter sequencing errors, and erasures. She stated that erasures signal a child who has an interest in improving his work. She cautioned, however, that some students focus more on correct papers than on learning.

Think Aloud and Observation Contexts. The videotape made for the think aloud consisted of children reading their "Ugly Duckling" stories to a duck puppet (handled by the teacher) who supplied assistance when needed. The teacher invited this researcher to observe a lesson that included a creative dramatic activity designed to stimulate oral language and oral retelling of a story previously read to the children. Students were observed to be actively engaged in the activity. However, as the teacher described words to elicit vocabulary from the children, they frequently gave incorrect responses. For example, in acting out the supposed funeral of Tom Sawyer, the teacher asked the children what you put the body in before you put it in the grave. After a student responded with box, the teacher accepted this response and said, "yes, but what do you call the box?" Eagerly a student yelled, "a body cask!"

Elaine's Assessment Profile

Introduction. Elaine's student population may be described as high SES and high ability. Her classroom was relaxed but orderly.
Students used their own homemade dictionaries for creative writing. Instruction did focus upon the higher order thinking skill of application. Instruction was teacher facilitated but students were also observed to direct part of their learning. The ability to perform independently was stated as a value by this teacher.

Table 3 records the frequency of data sources stated in all interviews for this subject. The total number of data sources identified was 260. A percentage of total responses is reported for each data category.
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Recognition</td>
<td>28</td>
<td>11%</td>
</tr>
<tr>
<td>Basal Test Data</td>
<td>21</td>
<td>8%</td>
</tr>
<tr>
<td>Comparison of Data</td>
<td>21</td>
<td>8%</td>
</tr>
<tr>
<td>Comprehension</td>
<td>20</td>
<td>8%</td>
</tr>
<tr>
<td>Behavior/Work Habits</td>
<td>20</td>
<td>8%</td>
</tr>
<tr>
<td>Work Samples</td>
<td>19</td>
<td>7%</td>
</tr>
<tr>
<td>Oral Responses</td>
<td>17</td>
<td>7%</td>
</tr>
<tr>
<td>Written Language</td>
<td>15</td>
<td>6%</td>
</tr>
<tr>
<td>Ability</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Oral Language</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Prior Educational History</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Motivation</td>
<td>10</td>
<td>4%</td>
</tr>
<tr>
<td>Intuition</td>
<td>8</td>
<td>3%</td>
</tr>
<tr>
<td>Willingness to Read</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td>Home Background</td>
<td>5</td>
<td>2%</td>
</tr>
<tr>
<td>Grade Level as Reference</td>
<td>5</td>
<td>2%</td>
</tr>
<tr>
<td>Gifted Status</td>
<td>5</td>
<td>2%</td>
</tr>
<tr>
<td>Observation</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Peer Coaching Data</td>
<td>4</td>
<td>2%</td>
</tr>
</tbody>
</table>

*Table continues*
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conference with Student</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Pace of Learning</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>ADHD</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Requests for Assistance</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Status as a Retainee</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Affective Response to Literature</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Birthdate</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td>Total number of data sources</td>
<td>260</td>
<td></td>
</tr>
</tbody>
</table>

Note. * = less than 1%. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Highest frequency data. Elaine identifies 260 sources of data in all interviews. Of these, 28 or 21% concern word recognition. This is surprising when one considers that the makeup of her present reading class is gifted and above average readers. Features of oral reading that provide focus for this teacher include fluency, number of errors, and the amount of assistance required. In addition, she reports recording words missed. She does not report any analysis of miscues. She states further that she has always relied on oral reading as a data source and considers it even more important for below grade level students. There is consistency between her self-reported strategies and number of actual frequencies in this area.

Sizing-up data. She extends the use of oral reading to a September sizing-up strategy. She summarizes her strategies: "I listen to them read orally... I have them write something for me then I do some silent reading and I do some questioning with them to see." Her responses are unique among subjects in this study in contrasting students' performance on oral vs. silent reading tasks. It is clear from her comments here that she is discussing comprehension independent of word recognition. Thus, she is not declaring that students find silent reading more difficult because they must independently decode. Instead, she uses the only reference to the concept of transfer found in all interviews in this study as she states:

A lot of times I can tell by their oral answers but that does not always show how they do on a written comprehension test... that sometimes is not the same... especially with a child who has problems... sometimes I think... they can sit down and discuss the story orally with you... but when they read a selection and
then have to transfer and read a question and find an answer. . . sometimes that is more difficult. . . and especially at the beginning of second. . . so that will tell me that's an area we need to work on.

It is also possible that she is referring to the concept of scaffolding whereby teachers provide support to students as they construct meaning in an instructional group. This support is not present when students read silently alone. Therefore, if students are unable to provide sufficient prior knowledge and vocabulary knowledge to construct meaning from the text provided, this may cause their silent comprehension performance to be depressed from the quality of their oral responses in the instructional group. Even students who comprehend the text may not be able to transfer the meaning they have constructed and apply it to the questions asked.

Independent work. Data from independent work samples are mentioned 19 times and are consistent with the statement:

...at these higher levels I think they need to be able to do something independently. . . you know, you pick up in the group on weaknesses and things. . . but when they can take. . . from the group. . . whatever you've introduced. . . or worked with. . . and be able to go and follow through with it. . . then I feel more secure about them. . . their being able to leave me. . . and go to somebody else and they wouldn't say, "what did she do with this child?"

The concept of transfer is again implied in this statement. In addition, her belief that it is important for students to perform independently appears to be a function of their perceived achievement level. Thus, more advanced readers should be more independent.
Another interesting feature is her concern over how her students will perform in the next school year and how that will reflect upon her as a professional. She states that she values oral responses in group work and their written work samples equally. Interestingly, the frequency of oral responses (17) suggests this teacher is accurate in her ability to self-report reliance upon data sources.

**Written language data.** In enumerating data collected for sizing up students in September, she adds written language (creative writing) to independent work, oral responses and oral reading. The frequency of written language statements (15) supports her stated reliance upon this source. When asked to talk about features of written language that relate to reading, she selects length, "depth," and "meat of the story." She conveys her expectation that a story will have a beginning, middle, and end. She feels that children who read more demonstrate wider word choices.

After many attempts to probe her perceptions about the use of students' written language for reading assessment information, she states that she views student growth in reading and writing to be parallel. She gains impressions of student ability and concept development from their creative writing. She summarizes:

...Writing over the years has become more and more...I love children's writing and I pickup a lot...what they understand from their writing. I guess I always go back to writing...because I think beginning readers...most of the time...good beginning readers to me are also good beginning writers...

However, she offers a warning regarding the use of writing to make inferences about a child's reading growth. She carefully
considers a child whose reading is fluent but whose writing may be unsatisfactory due to attentional difficulties:

... Something about the way they write that helps me know... but that's not a rule that you can put down in cement and never change it... because the one child who right now is in the gifted program that's in my reading class... his handwriting is terrible... and he has wonderful things to say and can't always get it down... he's ADHD... and on medication... but he's bright and he's interested in a lot of different things and he wants to put them down on paper... but he can't always... and sometimes doesn't want to... but if somebody who didn't know him were to look at most of the writing he does, they would not think he is a bright child...

Thus, this teacher identifies creative writing as a possibly invalid source of data for making inferences about reading proficiency for a specific subgroup of students. These are students who are verbal, impulsive, and distractible. They may or may not be strong readers, but are probably less developed in writing as a function of attention to task.

Comprehension data. Features of comprehension that provide focus for this respondent include details and main idea. She states:

... I know sometimes it seems that these tests try to zone in on a certain thing, but if I'm going to check comprehension, I want them to get a lot of things from the selection... I think to get a good idea of how a child comprehends, you check all those areas... 

Continued questioning in this area did not elicit additional
comprehension features. What is important to note, however, is the teacher's awareness of the purpose, content, and depth of coverage by the basal test. Nevertheless, she declares that she will measure more than that in the area of comprehension. Although there are numerous documentations throughout this study that teachers modify instruction because of externally mandated testing, this may be an exception.

Test data. In addition to her discussion about the lack of face validity of the criterion referenced reading test, she also expresses concerns in the area of reliability:

...Well, I do not care for the test... I do not think most of them tell you what the child really knows... like if you're checking comprehension there'll be four comprehension questions... if they miss one that's 75... which is a C... which looks like their comprehension is average... which it probably is not... so actually the reading test grades don't tell me a lot... I need more than four questions on a selection... I use it... but I use it wisely... I use it cautiously...

Later she states that eight or ten questions could provide useful information.

The reliability of data sources and the need to collect multiple sources of data to increase measurement confidence emerges as a common theme throughout all interviews. Elaine often alludes to the poor quality of basal tests. In fact, she has reviewed reading series prior to adoption and has had the opportunity to compare tests that accompanied eight different programs. She compares these and says:
I wasn't crazy about the Holt reading test (the previous adoption). . .but I don't like the HBJ much better. . .they're shorter and the Holt was just monstrous. . .slashes and computer cards. . .oh, we've been through it. . .I have not seen a series that I looked at the test and thought, "oh, this is a wonderful test" . . .I think maybe I'd like to write a reading test. . .because I haven't seen any that I really have liked.

She resolves this internal conflict of not liking the test but feeling she must rely upon some measure of vocabulary and comprehension, by discounting the sections she does not feel are valid measures or are unimportant skills. For example, in judging the merit of testing singular and plural possessives, she laughs as she says:

...I had a student teacher. . .and they didn't do well on that part of the test and she was really upset. . .I said, "honey, don't worry about it. . .how many adults do you know who use it wrong on their Christmas card or on their house?" To me, that's a dumb skill to be testing a second grade kid on!

In a similar manner, she chooses which workbook pages or black line masters to use with students by considering first whether she feels the skill is worth teaching and secondly, if the page measures it fairly.

Consistency of self-reports and data profiles. The only inconsistency in self-reported reliance upon data sources and actual frequencies for this respondent concerns basal test results. Although she states that work samples and grades are better indicators to share with parents than tests, the frequency of test
statements in this set of interviews is 21 (second highest category). This is easier to understand when one considers her statements about the use of basal test data. She appears to develop intuitions regarding student achievement and needs from interactive teaching; then she relies upon basal test data to confirm these impressions. In fact, she verbalizes that "waiting" is an assessment strategy. Her profile certainly supports that she compares data from a variety of sources. In discussing two students who changed instructional groups, she indicated that basal test results confirmed her observations in class:

. . No, the test scores weren't a surprise. . .no they confirmed. . .because they're weak in vocabulary. . .and their oral reading was not real strong. . .their silent comprehension was not real strong. . .in fact, I lingered longer on ___ than I really should have. . .because I knew he was really weak. . .but the next reading test he bombed. . .and I just. . .my instincts all along were that there was pressure in that group that he didn't need. . .

She stated that her strategy was to delay making grouping revisions based upon the first test score but to provide tutoring and reinforcement. She monitors and waits for the second test score. She comments on this deliberate waiting to make final judgment and says, "well, you have to work with a child for a while."

Observational data. Finally, in self-reporting her assessment style, this teacher describes how she arranges her classroom in order to teach small groups and observe students working independently at the same time. She quickly relates what might be considered behavioral data but gives her observations instructional meaning.
within the classroom assessment context:

...I tell children if you need a word...and I'm busy and you need help with the word...I say ask someone that you know will know it...so when you see a child...
you see somebody who's getting more than just one word or is asking this one and then turns to ask another one...
...and so you realize that they're having more difficulty than they ought to be having with something they're asked to do independently...

When watching a videotape of a classroom discussion she stated, "today he's turned around backwards when he is supposed to be reading or whatever, but yet he's grasping it all..." This suggests that although she notes his distractibility, she does not confuse his inattention with a lack of mastery of the content being presented. One would infer that this judgment is based upon prior experiences with this student.

Data collection and methodology. It is not clear how this teacher's orientation toward methodology affects her collection and valuing of data. She uses basal materials, workbooks, and basal tests. She also uses novels, journals, response to literature, writing conferences, and cooperative learning. Thus, it is difficult to characterize this teacher on a continuum between basals and whole language orientation.

Assessment concerns. What does emerge from the data, however, is that Elaine has some concerns regarding assessment in some whole language classrooms. She appears to view whole language as synonymous with little structure and a lack of comprehensive assessment of student learning. She reports that she formed this opinion after receiving several new students from other states.
Parents of these students characterized their first grade setting as a whole language program. Parents described their children as "doing fine in first grade." Initial assessment in their new school revealed that these children were still beginning readers. She summarizes her perception of the need for the documentation of data in whole language classrooms which specify a performance level.

...There's going to have to be some assessment... whether it goes into a reading folder or portfolio... whether it goes to an administrator... as a teacher, I still somewhere am going to assess formally or informally children and their reading... at this stage... because... I think with the whole language thing, you get all involved, and you get everybody involved, and everybody's doing things and everybody's cooperating... and somewhere... this little kid can't read... they're doing the things... they're following along... but they cannot read... and they come into second grade... non-readers... and nobody seemed aware of it... the parents didn't seem aware of it... the report card did not reflect it...

She felt this could have been avoided if the previous school had used a systematic assessment system. She concludes her statements on this subject with a concern that less experienced teachers will absorb the new philosophy and methodology, and miss the importance of a strong and accurate assessment system to document and report learning. This appears to be a reasonable concern for new teachers. Planning studies (Yinger, 1980) report that beginning teachers plan in chunks of content to be covered rather than in activities designed to reach a specified learning target.

Oral language data. Oral language appears to provide Elaine
with data about what words exist in students' expressive and receptive vocabulary. This is the same use of data found in Roberta's profile. This is interesting when one considers that the two respondents describe their students' abilities and needs as widely variant. Thus, it appears that oral language may provide as much useful information for this teacher of the gifted as for the teacher of remedial students. Elaine talks about an important distinction between decoding or word identification and knowledge of word meanings:

\[ \ldots \text{Sometimes you find the words are not always in their vocabulary.} \ldots \text{their everyday vocabulary.} \ldots \text{they can sometimes say a word but they don't know what it is.} \ldots \text{they really don't understand what it is.} \ldots \]

When asked if this was a greater problem with below grade level readers, she disagrees:

\[ \ldots \text{No, not necessarily because one little girl who's just a very good reader.} \ldots \text{and has her nose in a book all the time.} \ldots \text{she's always stopping me when I'm reading and saying, 'what's that word mean?'} \ldots \text{and if she doesn't know.} \ldots \text{I'm sure there will be many others who don't know.} \ldots \text{but she's alert enough that she wants to know now.} \ldots \]

Thus, this teacher can differentiate between instances of asking (or not knowing) word meanings as an indication of need or limited receptive vocabulary, and on the other hand, with the strength of monitoring one's comprehension during reading or listening. A very strong reader asks meanings of words because he or
she is unable to construct meaning and notices this instantly. Elaine describes activities she designs to measure vocabulary she has taught. She asks students to construct a sentence using the word or she designs a cloze activity.

Prior educational data. Prior educational data appear more important to Elaine than Roberta. While Roberta admits she may never read the entire cumulative folder during the year, Elaine reports that she checks them before the school year begins for birthdate, family situation, and evidence of frequent moves. She does not read previous report cards until she has graded the student at the end of the first marking period. Before she communicates with parents, she consults the previous report card to note inconsistencies. Elaine describes a triangulation strategy that will emerge as a common theme:

...After I do the report cards the first time, I go back and look...I don't want to make an evaluation of a child based upon what their past history is...but sometimes I look back and say, "am I having difficulty with this child and he's never had a problem with completing his work or never had a problem with reading. ...could this be me?"...and I'll look back...and I'm always greatly relieved...when I see that it is not a new problem that is surfacing...of course, I'm very happy if the problem wasn't there and still is not there.

Thus, she forms attributions from interactive data in instructional groups. These are verified by work samples and test data. Finally, she consults report cards to determine if this student's performance deviates from previous performance data.
Records from other schools and divisions are not helpful to this teacher because she finds the format confusing. However, narrative comments from previous teachers are important to her, particularly if she knows and respects the writer. She commented on notes from a previous principal she greatly respected. "Now, she really knew the children and if she wrote on the folder, I'd look at that."

**Grades.** Elaine has carefully considered what to grade. She does not grade workbooks, although she may choose a skill sheet she feels is a valid measure of what she has taught. She grades oral reading but does not inform students they are being graded. Further, she does not have criteria in mind for grading oral reading. She states, "it's more or less a judgment call." However, she does on occasion inform students of criteria for grading independent work such as, "today, I'm going to grade this for vocabulary." She explains to her students about averaging and the effect of not turning in a paper. Later, in referring to grading oral reading, she mentions "fluency" as important. She also grades spelling tests and basal tests.

Elaine feels she has some autonomy over what she records in her grade book. When asked, however, if she feels this autonomy means her grade book tells the true story, she responds:

"...well, if grades can tell a true story...I'm not sure anything in black and white actually tells a true story of what you're putting in a little mind...but as true as you're going to get!

However, she is realistic about the limitations of grades and when asked if the students who made C's learned less than those who had A's she responded, "Well, they're at least not able to demonstrate it."
Elaine stated she felt that the school division was likely to eliminate letter grades in second grade as they have in first grade. When asked to talk more about this and her concerns, she replied:

...well, I have mixed feelings about it...because I fought putting grades into second grade...because I thought we were pushing them down too low...and then you get used to working with them...you learn how to use them...and how to make them work for you...and now we're throwing it all out...and the report card that we saw had...what were those words..."proficient"...have you seen that?...the things are not going to be S and N or something simple...it's going to be "proficient," "developmental," or something like that...words that our parents at our school are not going to understand...and the parents many other places are not going to understand...and we can't figure out where they came from...we don't think they came from any teachers or anybody who works with children and parents...so the pendulum goes back...

This appears to record a resistance to change. Elaine did not want to change from descriptive grading to an ABC report card. Now she is reluctant to change back to a developmental scale. She supports her reluctance with the fear that parents will not understand what their children have learned. This fear may be also due to her experiences with parents of transfer students from programs described as whole language who did not know their children were beginning readers.

**Unique factors.** There is a unique factor which may affect Elaine's assessment context. However, it is not presented here as it
contains potentially identifiable information.

**Growth in expertise.** Elaine's responses to questions about her initial teaching experience and growth as an assessor in reading are candid and clearly describe the management issues of the beginning teacher:

...I had a really bad first year... I was a failure...
because I thought these are little children and if I'm good to them, they'll be good to me... they nearly killed me... it was awful... I had 27 or 28 kids and so... the inexperience was awful...

Informal peer mentoring was an important influence on Elaine's development as a teacher and assessor. She includes specific examples of the kinds of knowledge gained from mentors:

...Well, in that group of seven teachers, I was probably the only brand new one... So I relied on them for help... and I even a few times asked them to listen to children read to give me some ideas... also, I had not had any training in phonics... not as a student or as a teacher... and one of those teachers worked with me in phonics... 

In addition to the frequency of mentors mentioned throughout subjects' profiles, another common theme is introduced here:

...the importance of having mentors come into the new teacher's classroom and look at her students and her data...

Finally, Elaine states a very common sentiment among subjects...
in this study: they would not have continued teaching after their first year without the support and encouragement of mentors.

. . .I had nine children at the end of the year who couldn't read a word. Well, I thought it was my fault.. .that I was a failure. . .that I had done something wrong. . .I was too inexperienced to understand the readiness of the whole thing. . .and those experienced teachers persuaded me that it was not my fault. . .that these children were not ready to read. . .and had not had the experiences to read. . .

**Recommendations.** Her recommendations to improve the quality of assessments in reading include more emphasis on peer coaching. In addition, she has a novel suggestion for teacher preparation programs. When asked how we can teach preservice teachers how to assess their children and measure learning she replies:

. . .one thing I think William and Mary did this one year with their education students. . .they came on board the first week of school. . .and so they saw from the beginning. . .what an experienced teacher did to get a classroom under control. . .with management things. . .I'm sure they saw how a teacher would look at her groups. . .and the things that she would do that would assess them where they are. . .and I thought that was really a good idea for the teacher actually to see it from the bottom up. . .I think it would be helpful because in my own student teaching experience. . .I went in probably in October. . .well, she had everything all set as far as management, and as far as reading groups and everything. . .so I never saw. . .I saw what she was doing. . .but I
never saw how she came to that conclusion. . .

Elaine appears to refer to sizing-up strategies here. If this is an important phenomena of organizing a classroom for instruction, it appears reasonable that preservice teachers observe during this period of time and talk to experienced teachers about their thoughts about students as they get to know them.

**Written products provided.** Work samples provided were writing samples for three students (two samples per student) from September and February and a class creative writing booklet of prose and poetry.

**Think Aloud and Observation Contexts.** The videotaped segment chosen for the think aloud was a class discussion of *Pippi Longstocking*. Vocabulary was the focus of the lesson. The researcher was invited to observe a prewriting activity. Children discussed the word *adventure* and recalled Pippi's adventures. They were then asked to brainstorm modern settings where Pippi could have a new adventure. The classroom discussion of Pippi's adventures resulted in oral responses that were recorded on the board. Students were eager to respond and appeared confident. The teacher accepted all verbal responses. The lesson ended with independent writing. Children used dictionaries they had made to assist with spelling. Some children were observed to take their dictionaries to the teacher and ask her to enter a new word. No students were observed to have difficulty with the assignment and the last half of the observation consisted of the independent writing activity.

**Summary.** This subject is unique for her reliance upon paper-pencil data and independent work (29%). She appears very concerned over issues of accountability. This is reflected in her grading practices and consideration of previous report cards. One might
conclude that her assessment strategies are directed toward accountability rather than measurement of learning. She makes only one comment that reveals an attempt to understand how children construct meaning from text. In talking about bright children who do not know what a word means, she states, "they know instantly that it does not make sense to them."

However, when discussing writing she makes many writing statements that suggest she fully comprehends the process of constructing meaning. She looks for a beginning, middle and end. She monitors word choices. She may grade for vocabulary. She tells students what she is grading for. She believes the literacy activities of reading and writing develop together. She states specific examples when this does not occur and writing data may be an unreliable source of information about a student's overall literacy.

Reliability and validity are issues Elaine discusses in relation to basal tests. She uses results cautiously and weights sections she considers better measures of learning. She does grade oral reading, does not tell students, and admits she does not have criteria in mind. "It's more of a judgment call." This, combined with her emphasis upon oral reading with gifted students, suggest that she is not guided by a set of beliefs consistent with a constructivist view of literacy.

**Amy's Assessment Profile**

**Introduction.** Amy's school population could be described as high SES and low mobility. In addition, a high percentage of parent involvement is evident. Amy's classroom was quiet, orderly, and teacher directed. Students were on task with paper-pencil activities. Small group instructional groups were observed. The teacher could be characterized as nurturing and encouraging. Her interviews are replete with comments about her concerns for students'
self-concepts and how assessment affects students' self-perceptions.

Table 4 records the frequency of data sources stated in all interviews for this subject. The total number of data sources identified was 261. A percentage of total responses is reported for each data category.
Table 4

Data Frequencies for Amy and Percentages of Responses by Category

<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Recognition</td>
<td>32</td>
<td>12%</td>
</tr>
<tr>
<td>Comparison of Data</td>
<td>24</td>
<td>9%</td>
</tr>
<tr>
<td>Oral Language</td>
<td>23</td>
<td>9%</td>
</tr>
<tr>
<td>Motivation</td>
<td>22</td>
<td>8%</td>
</tr>
<tr>
<td>Behavior/Work Habits</td>
<td>18</td>
<td>7%</td>
</tr>
<tr>
<td>Oral Responses</td>
<td>18</td>
<td>7%</td>
</tr>
<tr>
<td>Comprehension</td>
<td>13</td>
<td>5%</td>
</tr>
<tr>
<td>Performance Assessments/Checklists</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Observation</td>
<td>10</td>
<td>4%</td>
</tr>
<tr>
<td>Ability</td>
<td>9</td>
<td>3%</td>
</tr>
<tr>
<td>Data from Reading Specialist</td>
<td>8</td>
<td>3%</td>
</tr>
<tr>
<td>Basal Test Data</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td>Work Samples</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td>Home Background</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td>Written Language</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td>Peer Coaching Data</td>
<td>6</td>
<td>2%</td>
</tr>
<tr>
<td>Prior Educational History</td>
<td>6</td>
<td>2%</td>
</tr>
<tr>
<td>Modifications in Testing</td>
<td>5</td>
<td>2%</td>
</tr>
<tr>
<td>Skills</td>
<td>5</td>
<td>2%</td>
</tr>
<tr>
<td>ADHD</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Intuition</td>
<td>3</td>
<td>1%</td>
</tr>
</tbody>
</table>

*Table continues*
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Level as Reference</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Willingness to Read</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Status as a Retainee</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Affective Response to Literature</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Birthdate</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td>Total number of data sources</td>
<td>261</td>
<td></td>
</tr>
</tbody>
</table>

Note. * = less than 1%
Highest frequency data. Amy identifies a total of 261 data sources. Of these 32 or 13% concern word recognition. Oral reading statements appear to occur most often in this data category. A unique theme that provides focus for this respondent is a student's willingness to read orally. She describes a reluctant reader: "and if I call on her...she's afraid...I listen..."

When referred back to this statement about a student's behavior in group when reading orally, she offers:

...when she's reading orally...you want me to tell you how I know...when she reads she gets emotional...and doesn't want to attack a word...so she'll show...emotion...she'll start to cry or she'll get real red in the face...so body language says a lot for her...so I'm trying to build up her self-confidence...by telling her it's O.K. to try...and her body language tells me a lot...and so I use that...

Specific oral reading features include fluency and expression:

...I see a big improvement and I know she's reading every night...fluency...and willing to work out the word whether it's in context or relating to the picture...A lot of times I'll put it on that progress report that they are having difficulty with their oral expression...

Other word recognition data sources include "word attack," and "knowledge of sounds." She explains:

...sometimes I have sounds that I think are real important...and if they're having trouble...I usually go by what the group is having difficulty with...
This is the first mention of checklists or any documentation strategy for performance assessment. In fact, Amy is unique among subjects in her utilization of these strategies.

Amy extends oral reading by including it as a September sizing-up strategy. She states, "...and then I listen to them read orally... and the vocabulary, and how they feel..." It appears that affective information is also an important source of data for forming initial expectancies.

In summary, there appears to be a consistency between the data frequencies for word recognition and this teacher's self-reported strategies.

Interactive data. Amy states she does not give many paper-pencil independent assignments and often handles guided practice orally. Indeed, her data profile lists six top sources of data that do not involve written products: word recognition, comparison of data, oral language, motivation, behavior/work habits, and oral responses. Therefore, it appears that over half of this teacher's data are derived from interactive teaching. The pattern of data here is also consistent with her self-reported strategies. She describes her teaching style and how it affects her data collection:

... I don't give a lot of pencil and paper practice. 
I give it orally... you know... like if we're talking... about compound words... I'll say, "Do you see any compound words in that paragraph we just read?" and I'll say, "What is a compound word? How do we know what to look for?"

When asked if there are ever any surprises when she assesses
individually, she states, "not for me. . .because I'm so observant. . . I can hear 'em. . . or you know, the way they respond." Thus, it appears that Amy has confidence in her attributions formed from interactive data. Her performance data provide a source of verification and documentation.

Performance assessment data. Amy refers to checklists, individual performance tasks such as providing a sound when given a picture stimulus, and notes from conferences with students, etc. As previously stated, she declares herself confident in her judgments; nevertheless, she values this type of observational and performance data, and documents more interactive data than all other subjects. What is most surprising is that Amy still states she is unsatisfied with the amount of written documentation:

... I probably don't have it down as well as I should on paper. . . you know. . . it's a lot of. . . you know. . . judgemental. . . I wouldn't say I'm real strong in that area right now to be truthful. . .

A logical inference here is that she understands the importance of documentation for observations or judgments, but cannot find time as a function of class size. In a later interview she comes back to this topic:

... I don't do a lot of it [performance assessment] because I find it so time consuming. . . as you can see. . . I haven't gotten very far for this nine weeks. . . and with 27. . . now if they go back down to where they're supposed to be. . . 20, 21. . . it's a lot easier than when you have 27. . . it's very difficult to do the one on one.
This subject is also unique among respondents in reporting a rubric for grading oral reading. Although she does not inform students of the criteria, or that they are being graded, she does consciously record these data as a performance assessment. She explains:

...Sometimes I'll have them come up and read to me orally...and I'll put oral reading [on the checklist]. ..and then I'll put check or check minus or dash if it's really poor...they don't know that...

Her assessment and documentation strategies appear to be responsive to parents' needs and demands. As a matter of fact, it may be inferred that the majority of the external mandates she describes and perceives are created by parents. For example, she states that she must have some documentation of oral reading because, although she might remember the performance accurately, the parents would be reluctant to accept this and label it judgmental:

...well, on certain things like oral reading...that's real important to me to remember 'cause how do you remember...you do know your children...you do know how they read...but sometimes the parents just say, "Well, that's just judgmental"...and you say, "Yes, but I've kept a record...and he read"...I'm very conscious of that...

Word recognition data. Although this teacher values word recognition data in self-reports and counted frequencies, there are no statements to document that she conducts any error analyses of responses. The following is a transcription with probes in [ ]. After several probes to identify the specific errors that caused her
judgment [the student is having trouble with sounds], she finally resorts to giving examples of behavioral data:

.. .let me tell you about N. She is having difficulty hearing sounds. ..[when did you first notice it?]. ..right from the beginning. .. so we've been working on that. ..[what would you see at the beginning of the year that would lead you to say that she can't hear sounds?]. ..well, she was very weak in word attack skills. ..and. ..um. ..[How do you assess that?]. ..well, in the reading group. ..and sometimes when we're doing other subjects. .. and if I call on her. ..she's afraid.

It is reasonable to conclude, therefore, that body language and emotionality during oral reading appear to be important features in informing Amy about a student's reading competence.

Oral language. In enumerating data sources for September sizing-up, she adds oral language and affective data to oral reading. Indeed, oral language accounts for nine percent of data (23 responses). This is the largest percentage for this data category among subjects. She appears to attend to vocabulary (word choices). In addition, she focuses upon information that is revealed through language about a student's background of experiences. When asked what their oral language tells her, she states, .. ."their background. .. what they've been exposed to. ..it makes a big difference. ..it makes a big difference." Like other respondents, she uses oral language to derive some estimate of ability. She then compares this estimate with the student's performance. For example, in discussing a student, she says:

.. .he has been labeled L.D. .. .he's a really smart
little boy. . .extremely smart. . .vocabulary. . .his verbal is outstanding. . .and so we're trying to emphasize that. . .a lot of times we have to do oral testing. . .he can't do any written. . .

During the year she reports tracking language growth by attending to length of utterance, ability to define words, and willingness to respond orally. She summarizes her students' growth in language:

... in a way. . .you can tell. . .I wish we could have done it at the beginning of the year and now because they would have answered in one word. . .but at least they were trying to expand their telling. . .I was really pleased. . .at the beginning. . .they would say one word. . .you know. . .they would say. . .in other words they would repeat the word back. . .report means to report. . .you know. . .uh huh. . .and that tells you that they're weak on their oral. . .so their answers. . .they're starting to expand. . .

Amy offers the only example of error analysis in the area of oral language. While watching the videotape and doing the think aloud activity, she offered the following analysis of a child's attempt to define the word brother. The first response in attempting to define the word was "a person." Next, another child said, "could be littler or could be bigger. . .could be older or younger." Amy added when watching the tape:

... I think he said "older or younger" because he was actually getting that from antonyms. . .old and young. . .because we've been doing that and I think maybe that
might have come across on this.

After he struggles to add to the meaning, the teacher said when watching the tape, "He's having a hard time with this concept...he did it [used the word to define the word] again." Thus, there is a high level of consistency between her characterization of her students' oral language and their videotaped responses. When additional probes requested her to describe what information a student's oral language reveals about their reading progress, she tied it to comprehension:

...If they can't express themselves, then I think they have difficulty understanding what the printed word is...because they really wouldn't know the context.

This subject was observed to accept all oral responses in both the videotaped classroom segment and in the actual classroom observation. Her strategy was to repeat their partial answers. This appeared to be an effort to elicit elaborated answers from students. She responded to a student's lack of response by repeating or restating the question. She was not observed to offer additional information or scaffolding to lead students from their partial understanding to a concept.

Observational data. This teacher links all oral responses with an additional component of interactive data, the accompanying behavior. The majority of Amy's behavioral statements are linked back to implications about a student's learning. In addition, transcriptions of Amy's interviews are inundated with statements regarding the effect that collection of data that will be valued to make judgments about students, has on students. For example, she frequently states that if a student is not successful in an activity
that is tested or graded, this will adversely affect their confidence and self-concept.

Therefore, the most essential characteristic of Amy’s assessment style is timing; she collects formal data only when she has enough interactive and performance data to tell her that students will be successful. If one totals the categories of observation and motivation with behavior and work habits, the aggregate becomes the chief source of data for this teacher. This is consistent with her self-reported style.

In further explaining her theory that collecting data affects students, she states:

...because I've been over the vocabulary. ... and the comprehension. ...I'm hoping that they'll be able to do. ...that. ...you know. ...without any difficulty. ...I wouldn't dare give it to them until I think they're ready. ...because it would be frustrating for them. ...and it would lower their self-esteem. ...and I'm big on self-esteem. I think that's real important. ...for them to do well. ...

Data collection and class makeup. This respondent does not produce any statements linking collection of data with class makeup.

Data collection and orientation toward methodology. Amy appears to exhibit an orientation toward a basal reading program and skills. In characterizing students' first grade reading background, she says:

...First grades used a lot of whole language and I'm not sure they had the skills. ...and the vocabulary. ...I felt they were a little weak from one of the classes. ...Two of the classes I would say supplemented their basal
with trade books and some whole language approaches.

...and this other teacher...just used whole language...

...and I'm afraid they missed quite a few very important skills...I felt...so...I went back and reviewed all those...and now...they're just really moving along.

This suggests that she favors a skill approach to reading supplemented with trade books. It is surprising, therefore, that Amy states that in her classroom, she uses very few workbook pages or dittos. When asked where she gets the majority of her grades, she replies:

...from worksheets...and then...not a lot...I make sure that it's assessing skills...it's not just workbook pages...just to get them to do it...it's after I've taught the skill...not in isolation but in context...

A logical inference from these two statements taken together is that Amy favors a skill approach to reading, but she teaches skills in a meaningful context. Indeed, her discussion of compound words supports this statement. In addition, she feels very strongly about integrating the literacy processes. "I think you have to have listening, speaking, writing, and reading and I think you have to integrate all of those together...and do a lot of writing."

Further evidence regarding Amy's orientation toward methodology can be gained from her instructional groupings. Along with reading groups, she teaches class groupings with novels, uses cooperative learning groupings, and uses peer reading partners. Further, in examining data frequencies, the percentage for basal test data for Amy is the lowest among all subjects.
In summary, Amy's stated orientation toward basals also includes integration of all literacy processes and an awareness of teaching skills in meaningful contexts. She could not be described, however, as a whole language teacher. This is interesting in light of the fact that her assessment style and data profile are more consistent with authentic assessment tenets. This leads to a conclusion that Amy's choice of instructional methodology and data collection strategies for assessment purposes are independent of one another.

**Written language data.** Writing is focused around the journal. She identifies a child's journal entry or creative writing as the best data to share with a parent to identify student strengths and needs. She selects invented spellings as a rich source of data. She talks about what information written language gives her about a student's reading progress:

> ...spelling. ...I think that's important. ...but in their journals I do not correct their spelling. ...it does give me information because they're sounding out the words. ...there again, they're attacking the word with sounds. ...it might not be exactly correct, but I can usually read every word. ...

Her actual behavior in analyzing student writing is not consistent with this self-report, however. There she chooses content and punctuation features and does not note spelling. She does however, appear to emphasize the concepts of sentences as whole thoughts and writing as conveying a coherent message. The actual journal entry is presented below in its original form:

> Today at school we had music. At music we song a song is was a lalaby song and we did a dance we followd
darectuns. We also had computer lab too we worked on time again and I know most of the time.

When asked what she would say to a parent when sharing this sample, she replied:

. . .I would say this child has a good understanding of putting sentences together. He's pretty much stayed to thoughts about today at school, the things he did. . .so his thoughts. . .he has organized his thoughts into a really. . .under his topic and I have really not dwelled a whole lot about staying on the topic. . .we talk about. . . but this child is. . .you know. . .is doing very well in organizing sentence structure. . .

When asked if any writing features had not yet developed, she stated:

. . .um. . .probably not with this child. . .this child is probably. . .pretty much. . .done what I have asked him to do. . .as far as. . .you know. . .sentence structure, capitalization, punctuation, complete thoughts. Once in a while you'll see that maybe he didn't stop when he should have stopped, but I think that's something that comes with more practice. . .

She admits later, however, that her satisfaction with this student's writing is also based upon her perception of his ability, "a good average student." When asked how she would change her expectations in writing for a gifted student, she adds:

. . .you would see more creativity probably. . .you know,
maybe more adjectives, which we have gone into. . .um. . .
little deeper thoughts that this. . .you know. . .like this child might have gone into. . .a child who would be a little higher would probably express their thoughts in not just sentences but, you know, say, "Wow! We had a great day today in school!" So, you know. . .using different words, structure. . .

Thus, it appears that she reacts first to features of word choice and language expression. In addition, she appears to hold different expectations for writing as a function of her perception of a student's ability. This is an "average student," and she says he has "done pretty much what I have asked." The ability to stay on a topic and organize one's thoughts are other important features of writing for several other teachers in this study who relate these to comprehension.

Finally, Amy relates writing as a data source for her overriding concern. She describes how writing provides information about student affect:

...I guess when I look at it I would say...again. . .is this great for this child. . .you know. . .my lower students started out just writing three or four words. . .and then they've really progressed to where. . .and I don't really correct. . .they share with the students if they want to. . .they're not made to. . .and I look to see if. . .actually, I look to see if they're having a good time. . .if they're enjoying it. . .you know, you can tell this. . .there's this little boy. . .you know, spelling not. . .and some of his sentences are not exactly correct. . .but. . .can you tell that that child
is really having a wonderful time. . . he feels good about it! . . .

**Test data.** Amy reports some unique methodology for incorporating basals into her reading program. These strategies include whole class use of a basal text, nonsequential use, double grouping students in more than one text, and use of basal materials to support a thematic unit involving speaking, reading and writing. She also reports giving the basal unit tests, although it should be noted again that her reliance upon this data source is the lowest among subjects. In addition, she reports that while she gives the results consideration, she does not adhere to the decision making parameters of the criterion scores. She states that she also considers the test results from the previous year and notes any narrative comments regarding modifications in testing, such as "read the test orally" or "retested."

She states that she does not formally retest but does continue her pattern of not testing until they are ready, using the test results as formative information, providing reteaching, and then informally rechecking for mastery. She describes her method for rechecking for mastery:

. . .I do go through mine [the basal test results]. .
. . .and I go back and give extra work and it's not really a retest. . . but we have more or less this year said we were going to keep moving them. . . because they would repeat the same skills or they'll be reviewing. . . so what I do is go back and reteach it through the reading group and then I'll give another worksheet or something to make sure that they have it and usually document that on my chart where I have main idea. I will mark that she
has mastered it. . .

Finally, she states that she compares a student's performance on similar content with different materials to insure that mastery has occurred. This system of repeated measuring and observing for the right time to administer the summative evaluation, provides this teacher with ongoing data for instructional fine tuning.

Thus, it appears that although Amy uses basals, she does so in a nontraditional manner. In addition, she incorporates many elements of a language based classroom: "I've really gotten on this bandwagon with the reading and writing and language across the curriculum. . .you know. . .like the journals, celebrity of the week, trade books." Her orientation toward methodology, moreover, does not appear to drive her data collection procedures. Rather, it appears to be her commitment to herself and her students that she will not test them until they are ready. She defines "ready" as able to perform successfully.

Grades. It is not clear why this informant presents some unique features in terms of collecting and valuing data and in using them in decision making. First of all, like other informants, she declares that she does not grade students' independent work that is an initial effort. However, this teacher extends this line of thinking considerably by stating that she usually does not assign independent work until she is ready to assess; she deliberately reserves paper-pencil tasks for this purpose. She explains how she gauges readiness for formal assessment:

. . .Oh, that's so important [oral language]. . .I think that's why I gotten away from all the paper and pencil. . .because I think that's really important. . .to express themselves. . .whether they're. . .it might be right. . .
. it might be wrong. . . but you never say it's wrong. . . you say, "Well, that was a good thought". . . you know, can you give me a little bit more. . . and you kinda pull it out of them.

When asked if she uses any paper-pencil tasks before the assessment she responds:

. . . Yes, I do that. . . I don't grade those. . . they're usually like. . . well, it depends. . . some of them are worksheets that come out of the regular workbook. . . which are skill sheets. . . and then some of them I've gotten from others or I've made my own by looking at other books. . . they'll do those. . . and then I'll go over them whole group. . . so that in the group situation. . . and then if it's a child that has. . . is having difficulty on that skill, or on those vocabulary words. . . then I'll pull them aside. . . and work with them individually. . .

The bulk of this teacher's formative assessments are conducted through analysis of oral responses, informal and formal performance assessments, group paper-pencil tasks that are checked in group to provide student feedback, and group assessment strategies that might be termed "checking for understanding" in teaching models such as Madeline Hunter. This might include "thumbs up," etc. Along this path of data gathering (none of which is graded until the final assessment), she intervenes with strategic reteaching in large group, small group, or individually. She describes her own process of change as an assessor:

. . . I don't give a lot. . . I don't give a lot of busy
work. . . I've cut that out. . . I used to. . . a lot of paper and pencil. . . but the last two years I've really cut back. . . if they maybe come up short on that [first paper-pencil assessment]. . . in other words, they don't do real well the first time, I ask them to go back and try again. . . and sometimes I'll even copy it over. . . I'll wait a couple of days and I'll say, "I really think that was just kinda a bad day. . . let's try it again". . . and a lot of times you'd be surprised. . . three days later they know it. . .

One could logically argue that massed short term practice followed by multiple assessments is not likely to produce a measurement of mastery that is stable over time. However, Amy is realistic enough to address this in her discussion of summative data as a function of content covered as well as time. She states:

. . . A lot of times if this little girl has to think on her own. . . for example, Freckle Juice was a novel that we did. . . and this was a cumulative grade of her retention of facts from the story daily. . . and when I added up, you can see it was a quite low score. . . and that's because she could not recall. . . so that would tell me that this child is still having difficulty. . .

Prior educational data. As with other informants, this teacher desires to shield herself from expectancy bias in September; therefore, she consults previous report cards after she has sized up her class. She talks about the information needs she has at the beginning of the school year:

. . . In the cumulative folder the only thing I really
look at. . .and I don't look at it at the beginning of the year. . .is the report card. . .I would rather just let the child come in and observe the child and see what that child is like. . .

Later she states she would go back and consult this information. There is no mention, however, as with several other subjects, that she would wait a marking period before looking at the report card. Indeed, her initial perceptions of her students and their needs appear to motivate her search for more information as she relates:

. . .Later I would go back to see. . .maybe I would go back just to see what kind of grades maybe the child was making. . .just to make sure I'm following up to make sure I'm doing everything I can. . .especially if it's a weak student. . .if I'm doing all the things that are necessary to observe that child. . .because I'd be concerned. . .am I seeing things. . .did anyone else see what I'm seeing?. . .

She makes a special note of one type of student for whom prior data would be especially valued, a student repeating that grade. Prior data on repeaters in the same school is identified as readily comparable and is valued because it is collected in a format and with rubrics known to the teacher. In addition, one infers that she painstakingly monitors children who are repeating a grade to make sure she makes the best use of this extra year:

. . .I actually knew of her [repeater] from last year. . . just being next door. . .you know, you know the children next door. . .but I did look at her report card.
I was interested in her report card. mainly. I guess because I wanted to see how much I could do to improve her from last year. I knew it was a critical year for her. not only self-esteem. but in her grades and so forth. so I did look to see her grades. I would say that maybe this folder would have more impact. this cum folder would have more impact for me than the other two. and that's because as I said, it is a critical year to help her really come out in her verbal or whatever she needed. and as you can see, she didn't understand. she didn't comprehend. she didn't use word attack skills. you set the goals for that child. the goals that that child really needs to make sure that they're on track.

Finally, another type of prior educational data is knowledge of participation in a special program, especially a remedial program. She thinks aloud about what that information means to her:

She did go to reading. I knew she was low because she went to the reading teacher. so I knew she was again scoring pretty low and needed that extra help.

Growth in expertise. In describing her growth as an assessor, Amy is typical in identifying mentors as an influence during her beginning teaching years. When asked what she focused on in evaluating as a beginning teacher, she admits to confusion and experimentation:

I think that it was trial and error a lot. when you begin. and as you progress in your growth and with
a lot of great people around you helping you... and if
you're willing to take the help... and not afraid of
constructive criticism... I think you have to be pretty
open to it if you're going to ask for it, you got to be
able to take it... and so I think after trial and error.
.. I think you go through experimenting... to see if
this works... or if this works better... so the first
part is really... you don't know what you're doing!...

She later offers that it was very hard to know who had learned
what was taught during her first few years. She further identifies
experienced teachers with 15 years of experience as the ones who
"really knew." She explains their influence:

... Yes... I think watching other teachers around you
that had had a lot of experience... when I started in
there were veteran teachers there... and so I would
listen to them... and watch them... in how they
perceive children... and work with them... like they
would have 15 or 20 years of experience when I started...
so I would watch them... I think it takes experience...

Amy identifies substitute teaching at a variety of grade levels
as having an effect upon her growth as an assessor of reading. She is
unique among subjects, however, in identifying professional
literature as an influence upon her assessment style. In addition,
although she is not the only informant to relate experiences with
their own children, she succinctly describes how parenting enhances a
teacher's ability as a "kid watcher":

... Um... when I wasn't teaching... I read quite a
bit... magazines... um... parents' magazines... and I
was a parent at that time. . .I would read those magazines. . .and it would tell you what to look for in your own child. . .so I would take that and try to look for those things in other children. . .

Further, she describes specific school-based staff development that provided strategies she now uses for checking for understanding:

. . .Staff development. . .I think that's real important. . .it depends upon what the staff development is at our school, we have a lot on assessing. . .group assessment. . .individual assessment. . .I learned to measure learning again through listening. . .in groups you can assess. . .individually and in group. . .you know, body language. . .like how many of you. . .if it's a group. . .how many of you. . .stand if you gave them a question. . .raise your hand. . .but a lot of time just body language. . .individually you can pull them aside and do hands on individual assessing. . .

Unique factors affecting assessment context. Finally, this teacher has eight interview statements regarding collaboration with the reading specialist. It appears that the influence from the specialist is not upon Amy's data collection strategies. Rather, it appears collaboration has influenced her style of assembling multiple pieces of data to profile and confirm or alter her attributions regarding a student's achievement or needs. She describes taking her grade book, work samples, notes from conferences with the student and conferences with the parents, checklists, test results etc., and then asking for the specialist's opinion. This is a substantial amount of student information to assemble and is consistent with her profile of
data frequencies. This teacher collects some unique sources of data and attends to diverse types of data sources in her classroom.

Finally, it should be noted that comparison of data is her second highest frequency and is confirmation of her style:

...I go to the reading specialist and I say, 'I'm having some problems with...how can you help me?'...if it is a student that she has, I say, 'What do you see?'....so a conference with her to see what she sees...are you seeing improvement?...or are you still seeing the same problems I'm seeing?..."

When faced with discrepant attributions, (i.e., parents' perceptions of a child's reading achievement does not match the school's), she seeks additional information to mediate this dispute. In this example, she identifies an Informal Reading Inventory (IRI) by the reading specialist as a rich source of information to provide additional data in order to reach consensus on a present level of performance. She retells a recent example:

...and sometimes she will test one of my students because the pressure is being put on by the parents, too. They'll come in and maybe they're giving you a little bit of a hard time here...and they'll say he's not making progress...and I'll get her to pull him...and I'll say, "How do you feel?"...I might feel differently from the parent and I'll say, "I think he's made a lot of growth"...and so I'll just get her to see what she thinks...

Peer coaching data. Perhaps one of the most interesting pieces of data reported in this profile concerns this teacher's use of peer
coaching data. She describes having her peer coach collect data on which children she calls on most, and analyzes this information to conclude she has called on her stronger students more. She also has instructed the peer coach to tally interruptions, their source, and her response in terms of teacher behavior. These data have provided information for professional reflection and growth for the teacher. As important, she is able to use it to learn more about her students as well as herself:

The peer coach comes in and evaluates whatever I ask her to evaluate. . .like if I say I would like for you to tell me as you watch me teach this lesson. . .am I calling on certain children more than others. . .like my low group or my high group. . .which is really a weakness that sometimes I have. . .I call on my faster kids. . .I do that. . .I'm better now. . .and so she can graph it. . .you know. . .you can have a seating chart. . .she can graph. . .you called on this one five times. . .you called on this kid one time. . .you could even have her come in. . .if it's discipline. . .I want her to tell me which children interrupt me more. . .or am I dwelling too much on whatever it is that distracted me?. . .

Thus, it appears that the above description indicates the teacher's awareness of the amount of data available in her classroom. The data collection strategies such as "she can graph it" reveal attention to capturing the classroom context accurately. Moreover, she declares that she needs assistance in order to attend to some of the interactive data (i.e., her own teaching behavior, etc.). Finally, she displays a unique openness to revision. In considering the teacher change literature, it appears that Amy is profiting from
staff development that has allowed her to confront her beliefs and change gradually over a period of several years (Borko, Flory and Cumbo, 1993).

Written products provided. Written documents provided by this teacher include a grid for recording individual performance assessment tasks and a home reporting form. The latter allows the teacher to rate content areas as well as motivation, daily work, homework, self-control, following directions, attitude, participation, attention, and cooperation. Ratings include "good" or "needs improvement" (See Appendix B). Work samples include creative writing and journal writing.

Think aloud and observation contexts. The setting chosen for the think aloud videotape was a reading group. Students participated in an oral activity to decode, define, and use new vocabulary in a sentence. During the viewing of the videotape, this teacher made several comments about students' affect and behavior. She described them as "jittery" and "nervous." She commented on their short attention span. She identified one child as needing more wait time and stated that she deliberately provides this. She also provided home background information on a child she perceived to be having difficulty. She accepted all oral responses with comments such as "almost." When asked about this, she stated that at the beginning of the year when she would call upon them they would "almost cry." Therefore, she is reluctant to indicate any oral response is incorrect.

This researcher was invited to observe a reading group also. A new story in the basal was being introduced with a lesson on vocabulary words. These were presented in context. Discussion focused on the word hero. When a student gave a partial answer, she repeated the response and asked for more. . . "not just help a person in trouble but something else." The next responses were "like
when you earn a medal," and "like when you earn some money." She asked the children to think of another word for "earn." She read the first two pages of text to the children and asked a student a question. When she received no response, she repeated the question and asked another child. On another occasion, she responded with "almost." She stated she does not record this type of activity.

In summary, it appears from these direct observations of classroom instruction that her concern for affect carries over into instructional methodology. That is, she is reluctant to indicate to a student that he is wrong. While she demonstrates some strategies, such as repeating a student's answer or asking for more detail to go with a partial concept in order to elicit more oral language, she was not observed to provide scaffolding deliberately in order to build oral language concepts.

**Betty's Assessment Profile**

**Introduction.** Betty's classroom could be characterized as structured, orderly, and teacher directed. Children listened for instructions and were on task. Whole group and small group activities were utilized. The decor of the classroom was clean and tidy as this teacher states she does not like clutter. Cooperative learning and small group instruction were observed. This school population is characterized by a high percentage of military dependents.

Table 5 records the frequency of data sources stated in all interviews for this subject. The total number of data sources identified was 193. A percentage of total responses is reported for each data category.
Table 5

Data Frequencies for Betty and Percentages of Responses by Category

<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation</td>
<td>26</td>
<td>13%</td>
</tr>
<tr>
<td>Comparison of Data</td>
<td>23</td>
<td>12%</td>
</tr>
<tr>
<td>Word Recognition</td>
<td>16</td>
<td>8%</td>
</tr>
<tr>
<td>Behavior/Work Habits</td>
<td>15</td>
<td>8%</td>
</tr>
<tr>
<td>Basal Test Data</td>
<td>14</td>
<td>7%</td>
</tr>
<tr>
<td>Data from Reading Specialist</td>
<td>10</td>
<td>5%</td>
</tr>
<tr>
<td>Oral Responses</td>
<td>9</td>
<td>5%</td>
</tr>
<tr>
<td>Home Background</td>
<td>8</td>
<td>4%</td>
</tr>
<tr>
<td>Grade Level as a Reference</td>
<td>8</td>
<td>4%</td>
</tr>
<tr>
<td>Prior Educational Data</td>
<td>7</td>
<td>4%</td>
</tr>
<tr>
<td>Oral Language</td>
<td>6</td>
<td>3%</td>
</tr>
<tr>
<td>Work Samples</td>
<td>5</td>
<td>3%</td>
</tr>
<tr>
<td>Letter Formation</td>
<td>5</td>
<td>3%</td>
</tr>
<tr>
<td>Comprehension</td>
<td>5</td>
<td>3%</td>
</tr>
<tr>
<td>L.D. Status</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Gifted Status</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Pace of Learning</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Status as Retainee</td>
<td>3</td>
<td>2%</td>
</tr>
<tr>
<td>ADHD</td>
<td>3</td>
<td>2%</td>
</tr>
<tr>
<td>Processing Skills</td>
<td>2</td>
<td>1%</td>
</tr>
</tbody>
</table>

Table continues
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modifications in Testing</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Standardized Test Data</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Child Study Data</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Interests</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Willingness to Read</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Peer Coaching Data</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Information from Parent</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Chapter I</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Hearing Impaired Status</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Data from Psychologist</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td><strong>Total number of data sources</strong></td>
<td><strong>193</strong></td>
<td></td>
</tr>
</tbody>
</table>
Highest frequency data. Betty's interviews contain 193 data sources. Of these 26 or 14% are observations. One is immediately aware that this teacher reports the largest number of observations among subjects as a data source for decision making. An investigation of the nature of these comments reveals that these observations occur during interactive teaching moments. Like Amy, inspection of the frequencies of subcategories of data reveals that approximately half of her information for classroom assessment is derived through interactive data. In addition to observation she reports behavior or work habits, oral responses, oral language, pace of learning, processing skills, modifications, interests, willingness to read, etc.

Thus, the teacher is kid watching as other students respond, read orally, proofread something on the board, or reread text to look for a specific answer. Betty is interpreting body language, facial expressions, movements, etc., to provide additional data regarding who has learned what has been taught. Examples of her interpretations during the think aloud include:

...M. likes to volunteer to read. ...even though she won't get it all right. S. likes to tell her every word before she has a chance to say it. The boy in the red shirt. ...if he's sure of the answer. ...he'll keep saying it over and over again. ...if he's not sure, he'll let the rest of the group have their say. ...D. is the one who tells everyone they're doing a good job. Now J.'s smiling. ...but she doesn't want me to call on her. ...she doesn't think she is as smart as the others. ...see her hands go wild. ...she is so excited. ...when she is excited the hands go. ...she is used to a small instructional group and she doesn't want to wait her turn.
to spill her guts... she is used to being able to say what she wants to say... when she wants to... and expand on the topic... P. can't stand to be wrong. He'd rather not give an answer... You see T.'s. hand going up and down just then... tap, tap, tap... he is very frustrated that I am messing with his sentence... 'cause he did not want me to rewrite it...

Although it is not possible to document the accuracy of these attributions, it is evident that this teacher has carefully noted specific behaviors during interactive teaching that provide data for reflection. Also, these counted frequencies are consistent with her self-reported style of reliance upon sources.

Betty's style includes reporting an observation such as "John is tapping his pencil." She then quickly follows this with a specific attribution for this behavior such as "he's bored with the pace of the lesson... or he just wishes we'd get on with it." In addition, her observations are characterized by specificity:

... M.'s writing is beautiful... she can copy anything but she has no idea what she is writing. You can't spell to M. and have her write it down... I spelled c-a-t for her and when she brought it up to me I said, "what is that?". She said,"cat, just how you spell it"... I said,"O.K." From then on, I would spell and somebody else in the room would write the word down for me and take it to her and show her what the letters look like...

Comparison of data. Betty also reports that she observes her students carefully in the classroom context and then tests them to confirm her perceptions. This may account for her large percentage...
of reliance upon comparisons. This was the highest percentage of reliance upon comparisons among subjects. Data frequencies, therefore, are consistent with her self-reports. She describes her thinking as she compares daily performance and test data:

...after taking two unit tests, they demonstrated on paper...what they had been demonstrating verbally in class...so that I had documentation to go before the reading teacher and say that they're as good in class as what these tests show; therefore, I want to move them up.

When asked to identify which students "really had" the skill after watching the videotaped lesson on cause and effect, she replied with confidence:

...who has this skill? T. has the skill down pat; J. has the skill down pat; P. has it; and the others fluctuate...R. doesn't have it...she can give it to me in piecemeal...if you tell her, "we're going to do cause and effect today...what can you tell me about cause and effect?"...she can do if she is guided...and break it down...if I just threw it up on the board...no...R. wouldn't be able to do it...

It is also interesting to note here that Betty differentiates levels of mastery among her students. Some appear to have guided mastery while others have independent mastery.

Word recognition data. This subject displays the lowest percentage of reliance upon word recognition among subjects. Nevertheless, it is still the third highest source of data for Betty with eight percent of data sources identified. She continues her style of specificity by clearly describing the learning target "knows
the words." These criteria are shared with students and parents. They are included in her remarks about testing vocabulary:

...Do they need to be tested on the words anymore?... don't know... all I know is I want my children to be able to know what the word means, use it in a sentence correctly, and be able to read it... and know when you see it what it means... so that when you get to another word... and you can't figure it out... you know the context of the sentence... and you're able to go from there... so on my vocabulary sheets that I send home, it says, "please make sure that your child can read this word and use it correctly in a sentence."

It is not clear, however, whether the sentence is to be oral or written. It is important, nevertheless, that "know vocabulary" means more than decoding words in isolation.

This informant's responses, however, suggest an equal valuing of word recognition and basal test data. Responses also reflect the valuing of word recognition vs. meaning-getting behaviors. In fact, several of her responses reflect her opinion that excellent fluency in oral reading in first grade may lead some first grade teachers to inflate their judgment of a student's overall reading achievement.
She reasons:

...I could see how a first grade teacher thought she was a good reader... she was... she was a very good reader... but not able now to give me what I wanted out of Level 7 at the beginning of second grade... but when we get to skills like paraphrasing... she just can't figure out how else to say it... if the author has said
it one way, leave it alone. . .that must be the right way. . .

This teacher attempts to weigh word recognition facility with an ability to perform abstract thinking which she sees as increasingly important in later grades. In discussing why she selected two boys for regrouping in order to provide more challenge, she responded:

. . .yes, some of the kids that are in the Level 6 group. . .they can read the words. . .which. . .the parents want them moved up to the seven group because they can read all these words. . .their decoding skills are great. . .however, when you ask them why. . .or what the author thought. . .or please predict, they couldn't do it they could only give you the set answers of what had already been done or what they had already experienced in their life where these two boys. . .were able to pick up the ball and go further with it. . .they weren't always right in what they projected into the story, but they brought more things into it. . .but they showed they were ready to go into the abstract. Everybody else was still on the concrete level. . .and these two boys were ready to go into the abstract which they going to need as they go on into the next level. . .that book is dealing a lot with author's point of view. . .so, I looked further down the road. . .

Thus, she clearly differentiates between parents' perception of reading as word recognition fluency and her view as getting meaning from text. Her comments suggest both a convergent belief system when she states "they can't give me what I want," and some
openness to divergence when she talks about a student "picking up the ball and going further." This is not the same as a constructivist stance, however. Betty is looking for a correct response rather than attempting to understand how her students are constructing their own meanings with the text.

**Performance assessments.** Performance testing of word recognition is accomplished through individual assessments with flash cards every two weeks (bottom group only). All documentation of numbers of words missed, error types, etc., relies upon memory. She explains:

.. No, it's usually just mental (record). . . I just do it mental. . . and they can tell you whether they got all their words right or whether they missed a lot. . . I'm not as concerned about how many they missed as much as I am giving positive praise when they get a word right when I know the last time they got it wrong I just keep it in my head. . . there's little drawers in the head. . . you know. . . like the Mickey Mouse Club. . . this is M.'s drawer. . . and this is A.'s drawer and this is S.'s and D.'s. . . and you know basically who can read what. . .

It appears here that she chunks information together by child. This assists her with recall of a student's performance. Later in discussing an increase in the amount of performance testing, she states the importance of having the teacher observe and record all performances. She feels that other professionals or paraprofessionals cannot communicate the entire performance through anecdotal records, checklists, or written documentation. She summarizes her concerns about performance testing:

.. performance testing. . . it takes more time. . . and
what do I do with my other 22 while I'm one on one with you? Some people say, "so get your aide to do it. . .get a parent volunteer to do it". . . If I get an aide or a parent volunteer to sit down and do my assessment, all I've still got is paper and pencil. If I do I don't have a clue as to whether the kid really did just make a mistake, did he get flustered, or does he really not know what is going on. . .cause I didn't see him. . .

Data gathering and class makeup. There is evidence that some data gathering strategies are a response to student need and therefore class makeup. For example, Betty relates that she collects performance assessment data on word recognition in isolation for her below grade level readers only. However, her many examples of observational data are not differentiated by perceptions of achievement. Furthermore, she appears as intent upon specifying learning targets, observing to make judgments, and testing to confirm perceptions with her gifted students as with disabled readers.

Data collection and orientation toward methodology. Betty reports her orientation toward methodology as "middle of the road" on the continuum between whole language and basals. She talks about her methodology:

. . .yes I find myself basically in the middle. . .I never bought into basal readers whole heartedly; . . .I don't find myself buying into whole language whole heartedly. I think it works for some children. . .I don't think it works for other children. . .um. . . .like organization. . .I like structure. . .whole language from what I've been able to view from people doing it. . .is. . .too much chaos for me. I look in the room and they've
got junk everywhere. . .and that bothers me. . .I do whole language in pulling everything together. . .but it was never given that term prior to last year such as science and social studies are pulled together. . .if I find a child that likes cars. . .then at reading time I read to them. . .I'll pull out a book on cars. . .

It is important to note that Betty's orientation toward instructional methodology does not appear to be in a state of transition to the degree of that observed in other subjects. In addition, she appears comfortable with her eclectic style. Teacher change literature identifies Betty's orientation as an "accommodator" (Borko, Flory and Cumbo, 1993). That is, she adheres to certain principles of methodology she has chosen and incorporates other elements as they are consistent with her program. In fact, her self-reports verify she is largely independent of methodological trends. She uses basal texts and tests, novels and trade books, performance assessments, reliance upon observation, videotaping of class segments, cooperative learning, and letter grades. While she values specificity of learning targets and consistency in assessment throughout the school division, she desires to have methodology within her decision making domain. She differentiates in her own mind between standardizing methodology and standardizing assessment in reading:

. . .but I also want to be left so I can be creative. . . I can be innovative. . .I can look at my whole class and say, "well, this is the way they need to get there". . .I don't want you to tell me how you want me to get there. . .but I do want you to tell me where you want me to go. . .besides, just be able to read. . .yes. . .that would
help me. . .

Assessment concerns. This subject clearly identifies word recognition as a concern as the school division moves toward a whole language program. Her data gathering behavior in response to her concern is to assess both basal vocabulary and vocabulary within trade books and novels that she teaches. She explains:

... it's a concern for me. . .the vocabulary. . .um. . .this next nine weeks, we're going to be doing *Ronald Morgan Goes to Bat* for the next two weeks. . .little book D. ordered for us. . .I went through the book. . .picked out words I thought should be vocabulary words at the same time. . .I pulled all the vocabulary words from Unit 2 out of *Weathervanes*. . .because I know. . .I know I should have been doing Unit 2. . .so they're getting all of those words. . .they're also getting words introduced in science and social studies. . .so that I'll know that at least they were exposed to them. . .

Betty feels that some standardization is desirable of testing materials, procedures for scoring, criteria for mastery, etc.

... it would be real nice if there was a chart or graph somewhere that said, "these books are reading level second grade. . .please pick any of these books to sit and have your children read". . .well, if I'm not going to use the reading book that has already been proven to be a second grade reading book. . .then I want somebody. . .reading specialist would be my choice to say, "here are ten books in the library. . .they're all written on the exact same level that your reading text was on. . .

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
when you go to test, why don't you use one of these*. . .
let them read a passage. . .then you can do your
comprehension. . .you can do your vocabulary. .
whatever. . .but it's strictly second grade. . .then
it's no longer the teacher going to the library. . .the
teacher spending her time trying to figure out which book
is appropriate. . .that takes away a lot of pressure for
me. . .you're the reading person. . .you tell me. . .I
can test 'em. . .but I don't want to spend my time going
out and looking for appropriate materials. . .

She argues that some standardization will help insure
accountability for certain teachers. She expands on this theme:

. . .well, again. . .I feel like good teachers are going
to be able to assess. . .I'm concerned as a parent when
my kid hits not a good teacher. . .I'm concerned. .
what's going to happen. . .nobody has all good teachers.

Finally, Betty is unwilling to accept teacher judgment in
selecting assessment materials. What is interesting is that Betty is
not concerned with invalid assessments resulting from teacher
selection of assessment materials. Instead, she addresses student
frustration:

. . .but there are a lot of books out there and it's
strictly teacher judgment. . .is this really a second
grade book? . . .is this maybe a little bit of a third
grade book but I think you guys can handle it. . .you
give it to them and they bomb. . .

Oral language. Oral language does not appear to be as
important for this subject. She reports this source of data less than all others except Stacy. Instead, she appears to substitute observation of body language, behavior, etc., for information other subjects derive from oral language. In addition, she displays such specificity in her ability to define learning targets that one might infer she does not attend to the quality of a student's language structure unless this is identified as an outcome in itself. However, another plausible explanation is that she processes oral language and its concomitant cues (prior knowledge, expressive vocabulary, concept development, background of experiences, etc.) automatically and never reports them in an interview.

**Prior educational data.** Prior educational data in this profile include psychometric information in confidential folders. Betty then compares these data to what she sees in the classroom context. It is as though she reads the diagnosis in the confidential, "visual memory deficit," and then calls up an image of that child in the classroom context. She selects the bits of data that help her to understand how that child processes visual information differently. She states:

. . .M. has a Child Study folder. . .a confidential. . .
and she has where. . .she has no visual recall. . .you
can show her the letter a all day today. . .and she can
tell you what it is. . .you come in tomorrow and you show
her an a and a b, and chances are good she has no idea
what letter she saw yesterday. . .

Knowledge of participation in special programs such as Learning Disabilities Resource Programs or Chapter I provides cues to this teacher about performance. In addition, she attends to a student's status as a retinee. This subject demonstrates the ability to comprehend prior data and relate them to two or three plausible
attributions which she then interprets within her classroom context. She illustrates this strategy while discussing S.:

. . .S. we're working through. . .he's already had an Ackenbach done on him. . .by both father and I. . .I just met with the psychologist and the L.D. teacher. . .and right now they can't decide whether it's a learning disability or whether it's behavior or whether it's home environment. . .or what it is. . .

Narrative notes on folders are valued prior data, especially those from the reading specialist. Betty displays the highest frequency of noting collaboration with the reading specialist:

. . .E. is. . .well, this is an interesting one. . .Eric is at the end of the Level 7. . .but I brought this because it has sticky notes from the reading specialist for you to see. . .

The reading specialist had written, "will probably need to repeat Level 7 next year." Thus, it appears that the reading specialist in this particular school participates in the decision making process of determining (or approving) mastery of material and placement in more difficult material for individual students. It follows that teachers would collaborate frequently with the individual responsible for this administrative task.

Finally, prior data include home background information. Betty addresses the importance of these data in understanding the whole child.

. . .oh, yes. . .D. . .beat the system last year. . .dad was remote. . .mother was working out of state. . .and he lived with grandparents. . .and he did no work at school.
. M.'s parents have been through counseling at different stages in their lives. . .but I suggested that they go back... 

Observational data. Work samples are only mentioned five times in all interviews for this informant. It may be inferred that Betty does not value paper-pencil products as much as she values her observations during work periods when these products are being completed. She again provides her interpretations of observed behaviors:

. . .I can tell by the way you sit. . .by the way you write your name. . .whether you are having a good day. . .and whether you are angry. . .

Moreover, she states that her annual project for her own reflection and evaluation is an ongoing video tape of classroom segments throughout the year. She feels this project will allow her to see growth as the year progresses. This is consistent with her reliance upon interactive (especially behavioral) data.

Test data. Testing plays an important role in confirming perceptions. She compares daily performance and test results:

. . .They're as good in class as what these test show. . .now comprehension starts to fall. . .he's still able to read the words. . .but he's not able to answer the questions about what's going on. . .then when we came up to the test, he scored one out of four in comprehension.

Although Betty gives all basal criterion referenced tests and consults with the reading specialist, she does not appear constrained by the decision making parameters of the criterion scores. Retesting
and modifications such as having a student read aloud or allowing the student to take the test in a distraction-free environment appear commonplace. However, all modifications in testing are recorded in the reading folder. She explains:

"...so we're going to let him go on and be exposed... however, at the same time, I go back and test him on Level Six. To see. Well, his test is perfect. He is making it up. He can't move at the pace of every four to six weeks doing a unit. Now this test was given by himself. And this [holding up another test] was given whole group. E. functions much better small group or by himself. That's why on here I had to document that he took the seven test with the whole group. It wasn't like this test where he took it by himself. Last year I had a child who could not pass a test if he had to read it silently to himself. He had to read it out loud. He had to hear himself. So once I figured that out. Then every test he took from then on was that way. And that's why E.'s test was given by himself. Now E. took unit three today and he didn't miss a thing on the test..."

Grades. Grades are computed, recorded and counted for a variety of purposes. Handwriting, for example, is graded for neatness. Grades are only put on the handwriting papers, however, as a motivator to encourage students to be neater. These are not averaged. It is apparent from Betty's comments that she understands the concept of using grades as a motivator (Frairy, 1992).

"...and I also put handwriting grades on my papers. I don't count those because there's no place on the report..."
card to show it. . .but I was getting all this slop city.

.soooo, I told them, "O.K. today, we're going to grade for handwriting," . . .and when I tell them we're going to grade for handwriting, all the slop goes away and the papers are written nice and neatly. . .and they take pride in what they are doing. . .so I'll grade for about a week. . .and then I won't announce I'm not grading. . .but they'll slowly go back into the slump and when they get back into there I say, "O.K., it's time to grade again". . .

There are also instances where grading appears to satisfy students' affective needs. For example, grades are put on workbook pages for the below grade level group even though they have already been checked and corrected in the instructional group. This is intended to help these students feel they are like the other students in the class. These grades are also not recorded.

Thus, it does appear that this teacher's data gathering strategies are sensitive to class makeup. In addition to collecting more word recognition data for students with perceived needs, she grades differentially and values grades as a function of her perception of student achievement in reading. She clearly does not average grades that are intended as rewards or motivators and are not reliable measures of achievement, such as worksheets completed with assistance in an instructional setting. Worksheets, board work, etc. may be accomplished with assistance from peers. While they are graded and recorded, these are not averaged. Thus, in a fashion similar to other informants, she does not count students' initial efforts with given content. When asked why she puts a grade on a paper if she does not record it, or why she records it if she does not average it, she replies:
I grade 'em because the kids come in expecting A's, B's, C's, and when they get a paper and it's just got a good written on it, or a check, they feel like they're still in first grade.

At least part of Betty's insistence upon grading (despite conditions under which the sample was collected) is in response to perceived parent needs. She discusses this concern:

...and parents give me a lot of flack because they don't know the difference between a good paper and an O.K. paper...but if there's six questions and I say you got five out of six and according to the system's grading scale, this is the grade you've got...then the parents are happy...the kids are tickled to death...and I know who sat and helped who...but I don't write that on your paper...

This teacher reserves a specific paper-pencil material for assessing a skill. She states that she does not assess until after she has taught it, reinforced it through games and worksheets, and perceives the students are ready to be evaluated. These are completed independently and monitored. They are graded, recorded and averaged. She discusses this process:

...the white sheets...Another Look...I use those for the grades after I've taught the skills...we do the games...we do the board work...we do the group activities...help thy neighbor...whatever you want...but when it comes down to time to do Another Look...that's strictly what you know...
Home environment and its effect on performance appears important in this teacher’s profile. She responds to this concern with a specific grading strategy. She reports marking in her grade book when a sibling is born, a father leaves for a duty assignment, etc., and then noting again when she perceives the family has stabilized after the change. In this way, she reminds herself to consider grades taken during this stressful period with great caution, and those outside that time as more representative of the child’s true performance level.

In communicating progress to parents, Betty feels letter grades are more informative than a more developmentally appropriate scale. She stated she was troubled by the proposed report card with DV and S. Expanding on this theme she said:

...but how can you tell the difference between a DV and an S...they told me it all comes down to a gut thing. ...but I can't accept this...and I'm not going to face parents and tell them that in my gut...I know your child is a DV in reading rather than an S...I feel you've got to have grades...I think documentation should be behind those grades...

These comments also suggest that Betty perceives measurement in reading, and therefore in grading, to be comparative rather than descriptive.

During parent conferences Betty shows parents the reading folder rather than her grade book. She feels the former is of greater value if parents can inspect actual items to gain an understand of what level of difficulty of reading was being measured as well as the format used for assessment. She explains information shared in parent conferences:
... when I meet with the parents. ... I always share the reading folder. ... I never share my grade book. ... and if the parents want to see the grade book, I tell them I will meet with them at a later date. ... and the reason for that is. ... that everybody else's grades are also in there. ... the parents get more out of the reading folder. ... where they can see the tested items. ... then they flip through the book. ... and they see the exact questioning the sentences. ... that's fair...

Finally, Betty recognizes that performance assessments yield more data about individual student performance and the context of that performance. Therefore, sharing this information with parents may take more time. When asked about the ability of others to understand a teacher's performance assessments, she states:

... that's where you have to sit down and you have to be able to say. ... why you did what you did. ... that's why with the parent conferences. ... I don't understand how you can do a conference like this. ... in 15 minutes on conference day. ... if you're actually going to sit down and talk to your parent about what's going on in an academic sense. ... how do you do it in 15 minutes? ... I want to know. ... and yet you want me to conference more with these parents to keep them abreast of what's going on. ... but I don't have any more time to do it. ... I can't teach for all the other stuff that they want me to do. ...
does state that she values parents' perception of their child's achievement level. This is unique among subjects. In most instances, a lack of consensus between parent and teacher on actual achievement level in reading is mediated by testing by the reading specialist. In this case the reading specialist provided the initial testing. However, the outcome is surprising and the teacher readily admits to a wrong decision based upon invalid data. Betty retells the actual sequence of events:

...when D. first came to us, he couldn't read any of the words. ...so I met with Dad and I explained. ...and Dad said he could read all of these words. ...he said, "do you have 'em?," and I said, "sure". ...so I went and pulled his reading test (an informal reading inventory), and I said, "now these are the words. ...this is the word that he said. ...and this is the word that he was supposed to read". ...Daddy flew hot. ...and said, "D., you can so read these words". ...Dad folded his arms. ...told D. to read. ...and he read all the way up to second grade. ...didn't miss a word. ...and I sat there with my mouth just hanging. ...he didn't hem and haw. ...just went straight down the list. ...bink. ...bink. ...bink. ...Dad said, "he can read. ...he just doesn't want to read. ...but he can read". ...

There is another unusual factor to note in this teacher's assessment context. Because of its proximity to a large military base, the student population is quite transient. This could also affect the frequency of collaboration with the reading specialist as children are continually tested and placed during the year. In addition, many second graders are attending their second or third
school. Therefore, this context may be unique because of the diverse educational programs students experienced in grades K and 1. Finally, parents in this setting are perceived to pressure children in areas of achievement and grades. This may significantly affect the assessment context. When asked about the most important thing she does as a teacher of reading, Betty states:

...I think my main thing is to take the pressure off of reading. ...they come in and they're so gung ho about these levels. ...levels, levels. ...what can I do? ...am I going to pass. ...why can't I be in the top group. ...and I just say, "everybody's going to learn all these vocabulary words. ...everybody's going to read all these stories". ...and try to take the stress out of it. ...
...when I meet with parents I also tell them. ...the only thing they're concerned about is levels so they can talk in the community. ...that's not good for their kids. ...
...doesn't matter whether they're in my top group, or my bottom group. ...

There may be some inconsistencies in Betty's response here and her grading practices in response to parent need. She stated earlier that parents prefer letter grades on papers in order that they can make judgments about whether it is a "good" paper or not. It is possible that placing letter grades on all student work that goes home exacerbates the amount of pressure parents place upon children.

Written Language. Betty's interviews do not provide any statements about students' creative writing as a source of data in assessing reading.

Influences upon assessment expertise. In tracing her development of assessment strategies during her career, this informant stresses her initial reliance upon paper-pencil data. As
she has developed, she has come to rely upon observational data. This is important also, in that this informant's description of her decision making data base is consistent with frequencies in her profile. Like the majority of teachers in this study, she underscores the need for mentors for new teachers. She feels mentors should come into the new teachers classroom and observe the dynamics of the classroom context. They would observe children, make judgments, and then interpret their perceptions to new teachers. One can infer that new teachers require scaffolding in order to receive this quantity of soft data that are floating about their classrooms: body language, behavior, oral language and responses, oral reading, social interactions, etc. With this assistance from mentors new teachers may be better able to construct meaning out of this mass of information to determine who has learned what has been taught.

At the conclusion of the set of interviews, each informant was asked if they had anything to add that had not already been discussed. This teacher provided the only response that added data to the study. She firmly stated:

...I think new teachers should be treated a lot kinder than what they are...they need a lot more support than what we give them. ...I think they expect to do it all well. ...sure they do. ...and if they don't, they're a fool. ...because you've got to project an image. ...you can do everything. ...your job's on the line. ...and I think we come across as saying, "do it". ...and then I think we veteran teachers come across as saying, "this is a piece of cake". ...because we make it look so easy. ...because a lot of the stuff we do. ...comes naturally now. ...after 19 years. ...but it didn't come after the first two years. ...and if I'm not willing to go down to
that new teacher's room and say, "I know what you see, but it's taken me 19 years to get here. . . and if you can do it in a year. . . there's something wrong with one of us. . . and I hope it's not me!!"

Written products provided. Papers provided included work samples for three children. The teacher identified one as an example of "far point copying." It could be characterized as lacking proper spacing and exhibiting emerging letter formation. Near point copying could be characterized, however, as displaying proper letter formation for single letter units and appropriate spacing and alignment. On a cloze activity for another student she noted erasures and stated the child was able to go back to the book and correct his errors. She commented on another student's paper on abbreviations by stating it had "many corrections." None of the papers provided were photocopied.

Think aloud and observation contexts. The classroom segment chosen for the think aloud was a lesson with the top group on cause and effect. It has been previously discussed. This researcher was invited to observe a whole group lesson on punctuation. This involved cooperative learning groups, board activities, and group written responses. While children did engage in reading and writing in order to proofread, it is not clear from interview data or field notes why this lesson was chosen as a good opportunity to assess students' progress.
Scotty's Assessment Profile

Introduction. Scotty's school population could be described as having many free and reduced lunch students. Moreover, students may live in rural areas. Scotty's classroom may be described as quiet, orderly and teacher directed. The testing session observed appeared to be distraction free. Students not being tested worked independently on paper-pencil activities. The teacher appears to focus many of her concerns around students' experiences, language, and home backgrounds. However, these data sources appear most important for her lowest achieving students.

Table 6 records the frequency of data sources stated in all interviews for this subject. The total number of data sources identified was 329. A percentage of total responses is reported for each data category.
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Recognition</td>
<td>37</td>
<td>11%</td>
</tr>
<tr>
<td>Behavior</td>
<td>36</td>
<td>11%</td>
</tr>
<tr>
<td>Comparison of Data</td>
<td>31</td>
<td>9%</td>
</tr>
<tr>
<td>Comprehension</td>
<td>28</td>
<td>9%</td>
</tr>
<tr>
<td>Oral Language</td>
<td>26</td>
<td>8%</td>
</tr>
<tr>
<td>Basal Test Data</td>
<td>20</td>
<td>6%</td>
</tr>
<tr>
<td>Prior Educational Data</td>
<td>19</td>
<td>6%</td>
</tr>
<tr>
<td>Oral Responses</td>
<td>17</td>
<td>5%</td>
</tr>
<tr>
<td>Written Language</td>
<td>14</td>
<td>4%</td>
</tr>
<tr>
<td>Work Samples</td>
<td>11</td>
<td>3%</td>
</tr>
<tr>
<td>Motivation</td>
<td>11</td>
<td>3%</td>
</tr>
<tr>
<td>Data from Reading Specialist</td>
<td>10</td>
<td>3%</td>
</tr>
<tr>
<td>Grade Level as Reference</td>
<td>10</td>
<td>3%</td>
</tr>
<tr>
<td>Home Background</td>
<td>9</td>
<td>2%</td>
</tr>
<tr>
<td>Intuition</td>
<td>9</td>
<td>2%</td>
</tr>
<tr>
<td>Willingness to Read</td>
<td>6</td>
<td>2%</td>
</tr>
<tr>
<td>Ability</td>
<td>6</td>
<td>2%</td>
</tr>
<tr>
<td>Skills</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>L.D.</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Conference with Student</td>
<td>2</td>
<td>*</td>
</tr>
</tbody>
</table>

Table continues
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Esteem</td>
<td>2</td>
<td>*</td>
</tr>
<tr>
<td>Standardized Testing</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td>Articulation Disorder</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td>Status as Retainee</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td>Total number of data sources</td>
<td>329</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* * = Less than 1%.
Scotty's interviews contain 329 identified data sources. This is the largest number among subjects. Kitty, who provided the least number of data sources in her interviews, identified only half as many frequencies of data as Scotty, in response to the same structure of interviews. Scotty's answers were more detailed. In addition, she felt comfortable adding related information.

**Highest frequency data.** Word recognition and behavior emerge as most crucial for Scotty. She displays much greater specificity than other subjects with regard to word recognition. While Elaine stated her students "could hardly read a word," and Amy addressed "retaining their vocabulary," Scotty draws very fine distinctions factoring word recognition evidence into her comparisons of data for a given student and finally formulating an attribution. For example during the think aloud, Scotty comments repeatedly that this activity appeared frustrating for this group.

...I did this activity a couple of years ago with children who were a half a year above grade level and they had a wonderful time with the activity. . .they loved it. . .just ate it up. . .they [this group] did not appreciate the activity. . .they were frustrated. . .it didn't make sense. . .they didn't seem to like it. . .it didn't keep their attention as well as some other activities I've done. . .the children I did this activity with before understood the parts of speech. . .these children had no idea at all about sentence structure. . .I think the activity was frustrating for them. . .

She then considers two alternatives and finally attributes their lack of success to their concept development. She rejects the notion that the readability of the activity was too high because she
relied upon student-generated vocabulary to build the activity. She makes a final comparison between this activity and others with student-generated vocabulary:

...the concepts...because it was words they themselves had generated...which they usually love that...before when we've used words that they've come up with and we've written them on the board and then we write sentences and stories, the activities have been more successful for most of them...and because they were using words like bobcat and animal words they usually liked, it should have been an interesting, more successful activity...yes, it was too difficult...and the fact that the sentences didn't make sense wasn't fun for them...the other group loved it when the sentences didn't make sense...this group...their frustration was really coming out...

Thus, Scotty confidently predicts whether a given student can decode a given text and conversely whether they could comprehend it.

This teacher focuses upon L., a boy who receives L.D. resource assistance. In considering his inability to read silently (or even to read to himself orally or subvocally) she recognizes that silent comprehension may be mediated by word recognition as well as distractibility:

...Right now L. is at the bottom of that group and he is not keeping up with the majority of the children in that group...I can tell from the fact that they are able to read independently...read a selection silently...or even if they need to read it orally very quietly...most of them are able to do that at this point. All of
them are making errors reading orally...all are...but they're able to grasp...get...comprehend what they're reading even with those errors...L. is not...and I haven't been able to zero in exactly what is keeping L. from learning...he is L.D. there are problems within the family...L.'s behavior is not...he does not use self-control...so there are a lot of factors going against L. being successful...

Thus, in the preceding sequence, Scotty retrieves and values information regarding home background, word recognition in oral and silent reading, comprehension, identified disabling condition, and distractibility. She realizes that an interaction is occurring to depress his comprehension whether reading is oral or silent because she states many other students make errors and are still able to derive meaning from the text.

Another fine distinction deals with the very different processes of comprehending oral language, directional vocabulary, sequential information, directions, etc. She further discusses L.:

...it's not just in reading...L.'s comprehension of even daily activities is not up with what the other children can do...uh...at times, he doesn't seem to understand the concept of just daily routine...and especially at the beginning of the year...to get the daily routine down was not easy at all for L...to understand where to keep his work, when to turn it in, and where to turn it in...to get that daily routine down...and now any time we do something new...and because he is so disruptive...it's hard for me to know
sometimes if he's acting out or if he doesn't understand or if he's acting out because he doesn't understand.

.or really see.

This suggests that she views his comprehension difficulties as possibly language related, although she does recognize that auditory memory is a function of attention.

Scotty identifies the Informal Reading Inventory as the most valuable source of information regarding word recognition along with observing students read orally. She discusses the IRI:

...P. and A. gave IRI's to those that were in our school last year so we would know exactly where they were the first week of school. ..I didn't understand how valuable a tool the IRI could be. ..

Like most subjects, she includes oral reading in her September sizing-up strategies. When asked how she develops her gut level feelings about students, she responds:

...I guess hearing them read orally. ..but hearing them read orally. ..how they do in group is where I learn the most from them. ..well, the first thing I do is listen to them read orally. ..one on one. ..I try to have the child sit with me while the others are doing an assignment or have the child stay in with me while the others are at recess or in a resource so that we're alone and so that I can hear the child read. ..

Although her responses in this area reveal a thoughtful attempt to understand the student's performance, Scotty's interviews do not contain responses regarding word recognition or oral reading criteria.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
or any documentation strategies.

Another very specific attribution is a comment regarding S. Scotty feels S. has improved in her reading but does not use phonic clues. She further explains:

...S.? yeah. ...well, she has just blossomed over the last couple of weeks. ...she seems to be retaining those words. ...her strength is sight reading. ...phonics. ...
...she's not as strong. ...that seems to. ...I don't know. ...
...if it confuses her or she doesn't differentiate the sounds. ...she's not really able to apply them when she sees a new word...

This teacher has kid watched sufficiently to describe the exact behavior S. displays when confronted with an unknown word. Scotty makes a very interesting attribution:

...She would just not respond. ...she would just sit and look at the word. ...and sometimes she moves her mouth. ...her lower jaw up and down. ...I think it's a nervous reaction...

Thus, Scotty selects a behavioral cue from Stephanie's performance to support her attribution. She responds to S.'s perceived nervousness with strategies designed to foster overlearning of sight vocabulary in order to build her confidence. An alternative explanation, however, is that this child may be subvocalizing, especially in consideration of her overreliance on sight word forms. She may be "trying out" the word she has decoded to see if it is a real word or if it fits the context. Nevertheless, sight vocabulary fluency is achieved and, in fact, may facilitate a deductive approach
to teaching phonic elements of known words.

Scotty's concern with word recognition as noted by data frequencies is consistent with her reported style. In addition it appears to be a response to class makeup. Scotty's class includes a group of students who are significantly below grade level. As with other subjects, the majority of her word recognition comments emanate from her concerns with this weaker group.

**Behavioral data.** Scotty makes 36 comments regarding behavioral data. Not only is this the highest number of behavioral comments and highest percentage of reliance upon behavior data (11%), she has twice as many comments about behavior as three of the other subjects. This is also regarded as a function of class makeup. Although this subject has two groups that are comparable in size (the other is above grade level), all but one of her interview responses for behavior are a product of her thinking about the lower group. Unlike Betty who attributes cognitive value to the behavior of any student in her class independent of achievement level, Scotty focuses her behavioral comments upon those who are not learning and attempts to understand their lack of success despite her efforts. She allows the interviewer to hear her compare data aloud:

> . . .they are able to contribute to the discussions and they're able to grasp the concepts and carry through. . .they're getting the main idea. . .of what they're discussing. . .so that gives me a clue that they do understand and know what to do and how to do it. . .and then again on work sample I'll see. . .that maybe on Tuesday when we did it they got it right and then on Friday when it's for a grade they rush through it and don't do as well. . .and it's the same skill. . .perhaps they know how to do it but they're not giving it their
Behavioral comments are always tied back to learning. In the one example of behavioral data with a very capable student, Scotty underscores the fact that this child's written work is not a valid measure of his achievement:

...I talked about C. in the top reading group...I felt his maturity...I felt his level of maturity kept him from doing his very best...and I still feel that this is true...so with C.'s work, because on the test...the level test at the end of each level...he usually gets 100% on these...everything...skills, vocabulary, comprehension, and then in group, when they are expected to do some seatwork...which the seatwork is very minimal this year...compared with...and C. just...his work is most of the time incomplete...or almost illegible...or he does the very least amount of work he has to...if it's a page where they are asked to write phrases or sentences...they are shortest he can give...and so the work samples don't tell me anything about his reading...when he can sit with the test and get 100%...that tells me he can read...because the tests are not that easy...especially for a child that's a half a grade level ahead...I usually think that perhaps it's maturity because with C., he has a difficult time staying with the group...when we're doing group work...keeping his attention, staying on task, even when it's in group...
The teacher's response to her concern (that she needs valid achievement data on C.) is to check orally for understanding in the group and to provide refocusing cues. She continues:

...I haven't found any way to motivate C. with the work...except that now what I'm trying to do...now that we're working with higher level thinking...if a child is immature, if they're developmentally not ready for that type of thinking skill, I'm doing more whole group activity and less independent work, because of children like C...that I feel they would lose a lot if I just said, "sit down and do it." I'm still having to say, "C., stay with us" a lot...I call on every child...whatever activity we're doing...I try to make sure that every child has responded in some way...and I really have to stay with C...keep eye contact with him...he wants to doodle...draw...and...

It is interesting at the conclusion of this discussion to hear her appraisal of the success of her chosen strategy. When asked whether oral responses tell her more she responds:

...well, it still tells me that I haven't motivated C. to...I was hoping that it would be of greater interest to him with the oral...with children throwing out ideas...discussion type activities...with even less seatwork...so far it's not working quite as well as I'd hoped...but I'm hoping in time...it's been a month...so I'm hoping that as time goes on he'll...it will draw his interest a little more...
Comparison of data. Scotty's decision making style is clearly dependent upon comparing data from different sources. This accounts for over nine percent or 31 of her total responses. She verbalizes in the previous excerpt how she compares a student's written work with his oral performance in group. Later she compares oral language to written language responses with remedial students. In comprehension she contrasts oral and written responses. This was also observed in her previous discussion of C. She appears to be validating her own judgments by collecting various data on a given student and looking for consistency. She identifies discrepant data as a possible sign of invalid measurements. For example, when she discusses C., she characterizes him as a bright child with excellent oral language, concept development, and background of experiences who performs well in group and on criterion referenced tests, where he does not have to generate a response but only select an answer. She then contrasts this success to his independent work samples or creative writing where his incomplete work, or immature and brief responses might lead one to conclude that he lacks the concepts, or skills, or both. She knows that this is not the case, and points to his young age, distractibility, and home factors as variables.

Although she readily admits to comparing multiple data sources on a given student, Scotty denies looking for patterns over period of time. She also rejects the notion of reducing the complexity of decision making over the years by developing consistent conclusions from similar data sets. She firmly states that there can be many causes for a set of data and that she looks at each child individually. This statement is challenged when the nonsense grammar activity for the think aloud is not successful. As Scotty looks for causes to explain students' lack of response to this activity, her first thought concerned data gathered on another set of students.
years before when she had taught the lesson. Thus, it does appear that Scotty chunks some information in dealing with the simultaneity of data in her classroom. It is not clear from her responses whether there is consistency in her self-reports at the individual pupil level.

**Written language.** Her reliance upon written language is the second highest among subjects. Indeed her self-reports and data frequencies are congruent in this area. She states confidence in her ability to make judgments about a child's reading development from this data source. First of all, she assesses students' knowledge of vocabulary from their ability to generate their own sentences with the word. Further, she explains her criteria in judging this work:

...the children had to use vocabulary words that they had spent quite a bit of time on in sentences. This is his own writing and spelling. I would look first of all, for content...whether or not the child understands the meaning of the word...and how to use it properly in a sentence...checking somewhat for grammar...but since we don't spend very much time at all on grammar per se, I would correct it and hopefully, as we use words in group they would start getting a better grasp on grammar...

Secondly, it appears her methodology in reading, especially with the lower group, is inextricably bound to writing. She accurately defines herself as a language-experience teacher:

...I'm very much interweaving the vocabulary...I'm going very strongly with the vocabulary from the basal but I'm using it in activities...I'm using language experience a lot...the children are dictating a lot of
stories where they are incorporating the vocabulary into those stories. . .activities that we're able to incorporate the vocabulary words. . .for instance we went on a walk outside. . .right outside. . .on the school grounds. . .and the words that we were working with were playground, swings, and so I sort of pick and choose. . .I look at the vocabulary words and I base an activity that those words will fit in. . .then the children. . .when we came back to the classroom. . .we talked about our walk. . .and then they wrote a story in group about the walk and I had the words listed on the board and they were to draw from those words to create their story. . .

This appears consistent with the definition of language-experience classrooms such as those based upon Russell Stauffer's theory (1980). In these classrooms, literacy learning builds on children's experiences and capitalizes on the strength of the oral language. Thus, reading is talk written down. While this approach has many commonalities with the constructivist movement, it also has fundamental differences. The greatest of these is the concept that reading is a meaning that is constructed by each reader from the text provided. While it is mediated by the child's experiences and oral language, it is more than "talk written down."

Finally, Scotty includes writing in her September sizing-up activities in order to derive cues about students' thinking skills and comprehension of language:

. . .I also look at their ability in writing. . .we start writing the first day of school. . .and how much they're able to write. . .if they're able to write a page. . .if
they're only able to write one sentence. . .some of them cannot even write one sentence. . .they have to draw a picture. . .and that gives me an idea of. . .a very good idea of where they are. . .in their thought processes, in their vocabulary. . .

Oral language data. Just as reading and writing appear intermingled, so do clear distinctions between writing and oral language. She states:

. . .if they wrote five sentences. . .and only two of those made sense. . .could be comprehended. . .then I knew that that child needed remediation on that set of words. . .and more application in group. . .being able to experience using that set of words in a story or even in just talking. . .where there's no writing involved. . .and so I knew that that child needed that language. . .

The most important tenet to note here is that Scotty appears to feel as a teacher of reading that she is also a teacher of oral language. She appears to demonstrate an understanding of the concept that oral language underlies all literacy learning. Her percentage of oral language responses is the highest among subjects. Scotty's thinking appears transparent as she walks backwards from speaking to comprehending to writing. All of this occurs within a sea of experiences as she explains:

. . .if they could not form a sentence that made sense. . .one time I asked a little boy. . .we were working on a set of words that had to do with forecasting the weather and predicting. . .and it was a story about forecasting. . .and I asked him to just look out the window and just
tell me what the weather was like today. . .and he
couldn't do it. . .he could not form a sentence. . .he
would say, "sun". . .or he would just give one word
answers. . .if L. has to write a paragraph about
something like a cat. . .he might be able to write: "The
cat is black. The cat sees a mouse. The cat eats the
mouse." But if he had to do a creative writing activity
like I was talking about where they had to write a scary
story. . .and we were using the topic of dinosaurs. .
.which I knew they loved. . .and we had been reading
stories about dinosaurs. . .so most of them. . .almost
all of them. . .were able to generate good ideas. . .L.
was not. . .if he had to do a creative writing activity
where he had to pull from experiences that perhaps he has
not had or didn't comprehend, then he wouldn't be able to
do that. . .but if it's something that he has had lots of
experiences with. . .like a cat and a mouse. . .he could
do a basic paragraph on that. . .

Thus, the two factors in predicting L.'s ability to write on a
given topic are his ability to form primitive sentences orally on the
topic by drawing upon his prior knowledge, and whether the
manipulation of language falls into his experiential background. In
addition, the preceding thought sequence illustrates the concept of
teacher prediction: teachers regularly manage the complexity of the
classroom by their ability to predict students' success on given
activities (Yinger, 1980).

Like other subjects, Scotty uses oral language for cues about
general ability and compares this with a student's performance for discrepancies (See discussion of C.). When students display limited oral language, she collects more language data. She finally concludes L.'s reading problem is not isolated to this area and she must take a more global approach to assist him. This decision is informed by her observation of his ability to comprehend everyday social situations at school and follow directions. She sums up:

...and it's not just in reading...but pretty much across the board in school...L.'s comprehension of even daily activities is not up with what the other children are able to do...uh...At times, he doesn't seem to understand the concept of just daily routine...and especially at the beginning of the year...to get the daily routine down was not easy at all for L...to understand where to keep his work, when to turn it in, where to turn it in...and now any time we do something new...and because he is so disruptive...it's hard for me to know sometimes if he's acting out or if he doesn't understand or if he's acting out because he doesn't understand...or to really see...

What is interesting in this discussion of L. is that, although L. is identified as L.D., Scotty does not relate information about his processing skills or whether he also receives language therapy. Finally, this experienced teacher verbalizes the process of guided oral practice. In this case independent performance is delayed for a month to allow these students to achieve oral competence in forming a story. Scotty explains the process:

...we incorporate the vocabulary into the experience and that time it was our second activity and second
story. . .we worked as a group on it. . .and we continued
for about four weeks doing only group stories. . .they
could hear each other's thoughts and ideas and sentences
and words. . .Then I began to ask them to generate their
own stories on their own independently. . .but not
actually writing yet. . .they dictated. . .and they are
to use the topic. . .

Scotty demonstrates an understanding of moving students
gradually in acquiring language concepts and applying them. The
resulting group-dictated and individually-dictated books are
documentation of students' oral language.

Prior educational data. Prior educational data consisted of 19
responses for six percent of total responses. These data for Scotty
include a student's status as a retainee, L.D. label or Chapter I
student. These labels appear to provide cues about achievement.
Home background appears very important to this teacher. Again, this
occurrence is viewed as a response to class makeup. She provides an
eample of the type of information she considers in explaining a
student's lack of progress:

. . .there's really one in particular that I feel is
really floundering. . .this is a learning disabled child
and because of things that are happening at home. . .with
his parents. . .they are in the middle of a custody
battle over him. . .and he acts out because. . .we think
because of this. . .he's the one that right now. . .we
haven't seen a tremendous amount of growth. . .

Prior information also includes cumulative folder data. This
subject is unique in the amount of attention she focuses in this area. She lists the types of information gleaned from the folder:

. . .I always look at the parents' occupation to see if their parents are working outside the home. . .which give me an idea of perhaps how hectic their schedules might be. . .although I have found that even when one is not working outside the home. . .their schedules can be quite hectic. . .I look to see how much education the parents have. . .so that I might weigh the experiences the child will get outside of school. . .how much general knowledge that child might have coming in to me. . .or get throughout the school year working with the child. . .I check to see if it's an occupation where I might be able to invite the parent in. . .to visit the classroom. . .which would add to our curriculum. . .lots of good information there. . .just knowing what the parents do. . .and how many siblings there are. . .knowing the age of the siblings. . .if the child is the oldest, the youngest, the middle. . .I would see what teacher they had. . .if I didn't already know that. . .

Like other subjects, she looks at previous report cards. However, she states she would consult these immediately in the fall:

. . .I look at the report card. . .not so much usually to see what grades they have. . .because it's usually from first grade. . .its very general. . .satisfactory, very goods. . .which is sometimes very hard to tell. . .exactly what went on. . .but I many times look for
behavioral problems. . .if I'm experiencing behavioral problems with that child. . .to see if that happened the year before. . .to check to see how that child did socially. . .I try to glance at them when I'm removing files and folders from the cumulative folder at the beginning of the school year. . .I try to glance at the report card and see overall how that child did. . .and then if that child begins to experience difficulties in any area. . .academic or behavioral or social. . .I go back and look again. . .

This comment also reveals that the lack of consistency in the grading scale from first to second grade, makes the information less usable. One infers that she views letter grades as more specific. Also like other teachers, Scotty values narrative comments from teachers. However, like Elaine, she believes that these notes make teachers libel:

. . .even though we're not allowed to write notes like that. . .but if there's a modification, we can note that. . .but I would want to see. . .but I don't think we're allowed to write notes in the reading folder. . .we can't put personal opinion. . .but if there's modifications. . .Thus, evidence of testing modifications are valued prior data:

. . .I do look to see what level or I might glance a the last test to see how they scored. . .if they just did average or above average. . .were there any modifications. . .because I always note modifications at the time on the cover sheet when I'm testing children. . .hopefully, other teachers do. . .
Finally, previous IRI's are only valued if they are recent. When asked whether reading testing in the reading folder would be valuable, Scotty replies:

. . .it would depend upon when it was done. . .if the child is coming. . .say in September. . .and the IRI was given in May. . .then I would give a great deal of weight to it. . .I might want to ask to see it repeated to see if the child lost or retained over the summer. . .but if it was given at the beginning of first grade, I really wouldn't give much weight to it. . .

Work samples. This teacher appears to consider carefully the limitations of using work samples as data for decision making. She is well aware that word recognition, distractibility, and maturity are variables in the quality of written work. Any conclusions, therefore, are tentative unless supported by other performance indicators. She states:

. . .because a lot of time a child just does not do the seatwork. . .or they're not on task enough to complete enough of the seatwork. . .to really see where their capabilities lie. . .

She acknowledges that her own growth and many years of experience have made her aware of the lack of reliability of work sample data. Over the years she has learned how many mediating factors may affect this data source:

. . .I think it was in seeing that there were just some children that were never on task. . .were not ever able
to complete work. . .and would sit all day and not do any
work. . .and no matter what I tried, they still were not
on task. . .that seatwork is not as reliable a tool as
what they're able to do when I'm working with them in
group. . .

Thus, it appears that as Scotty distilled wisdom from her years
of teaching, she also shifted the source of data to inform her
decisions in reading, from work samples to interactive data in
instructional groups.

Test data. Test data are viewed as one more piece of
information. This teacher states she is comfortable modifying tests
to allow for a distraction free administration, or to allow a child
to read to himself orally. She explains what she feels constitute
appropriate testing modifications:

. . .I've learned how to use modifications for the test
so that I still feel they're accurate and reflect how the
child can read. . .but I've learned how to modify the
test and meet the child's needs. . .especially their test
taking needs. . .if they need to read orally. . .if they
need to stay in during recess and read when no one else
is in the classroom if they are easily highly
distractible. . .if they need to read a selection twice.
. .

What is most interesting in the area of test data are the
differences between self-reported strategies and those actually
observed. Scotty reports that she does not assist students with
words:
I never actually read words for them...I don't give them clues like "it rhymes with train" or anything like that...I try to give them test taking strategies...so that if they're taking the test two or three years down the road and cannot ask the monitor any questions, they'll be able to cope...

In addition, she states the importance of documenting any nonstandard testing accommodations in order that the data can be accurately interpreted. However, in reality some assistance is provided. When a testing session was observed a child went to the teacher and asked, "What's that word?" The teacher responded, "What is Newport News?" The child said nothing and the teacher repeated the clue. The child then said the correct word which was city. The researcher observed seven children go to the teacher for word recognition assistance. One child came five times. In assisting with the word celebrate, the teacher said, "What do you do on your birthday?" The child responded, calendar. The teacher said, "No, we have a calendar up all year. You invite people over to help you_____." The child said the word correctly. In the final example chosen the child came over and said, "I forgot what this word is." The teacher covered the ed on dream. The child responded, "dr/e/." The teacher said, "that's the way it begins. What is the vowel?" The child then responded with the long e sound. The teacher said, "O.K. What's the word?" The child responded correctly. It is not clear whether this test was given diagnostically and therefore in a nonstandard manner. The teacher remarked to the researcher that she only gave this test in case the teacher next year wanted this information. Her conclusion after the testing session (before the tests were graded) was that she felt the test was too long and difficult for this group. It was her recommendation that this group
have continuing instructional experiences with concepts and vocabulary tested here. Therefore, she felt they would need review and direct instruction at that level the following school year.

**Grades.** Grading appears dependent upon a number of noncognitive factors including effort and Scotty's assessment of general ability. She describes grading a student who had great difficulty formulating sentences:

...like one time we were trying to write a story about something scary...and this child that couldn't verbalize about the weather...his story made no sense at all...he had written some sentences down but they didn't...it just made no sense at all...so...if I remember correctly, I don't think I even graded that story... 

When asked how she grades students who try but cannot begin to perform on second grade tasks, she replies:

...well, say if they wrote five sentences and only two of those made sense...could be comprehended...I would just weigh that child's abilities and I might give that child a C because I felt that was average for that child...when I would grade that child's writing...I would not say, "well, Sue over here was able to write a nice little paragraph. She makes an A; this child going to make an F because he cannot write a nice little paragraph...I would look at what he did write...and assess how successful he was with what he put down on paper...sometimes I would give a D if I felt that the
three sentences they generated were not sentences and
that they did have the ability to write a sentence...

Communicating progress to parents with letter grades also
reveals inconsistencies. When asked what she would share with
parents, she relates an actual sequence of events with a student
below grade level who was making A's on her report card. Scotty was
asked to select among three sources of data to share with parents
toward the end of the school year: her grade book and daily work
grades, the reading folder and testing, or the cumulative folder.
Her response is the narration of a conference with the parents of S.
on the last day of school:

...well, I had a conference with the parents. Their
child was in this reading group. They first wanted to
look at the report card...and they wanted to know about
the reading...level...and so we talked about that.
Now this child came in to me as a preprimer reader and
left on a second grade level. But mostly we looked at
the IRI, because that's where I put the most weight...
.as to truly judging what she could do...and so that
the parents could understand...and we talked about what
they could do over the summer...because they very much
want her to be reading at what they call "grade level"...
.and I explained to them about us moving into whole
language...and what S. perhaps would experience in
third grade with whole language...but we mostly talked
about the IRI. She had all A's on her report card.
Anything that I asked her to do, she was able to do well.
She did very well on the comprehension and vocabulary

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
tests. She was doing very well with the work that we did in group. S.'s parents questioned how she could have an A when she was below grade level...and we talked about the fact that because she is doing outstanding work where she is...you don't give someone an F because they're not reading in a certain book. I always try to make the parents understand...if nothing else, we send interims home. We do send interims and it's marked on the report by code that the child is below grade level...but I usually especially if I think the parents will not...that they might have misconceptions about their child's ability because of the letter grade...and that's another reason that I'm glad we're going away from the letter grades...because of the misconceptions that some parents have...that their child has an A so they're doing just great...and that there are no problems.

Finally, she appears to concede that the mix of effort, achievement, and ability is not the same for all students:

There is no happy answer because we just don't grade all the children the same way...that isn't really the right way to say it...it's because children are not reading at the same level...their grade can be deceiving...

Unique factors affecting the assessment context. There are two important factors that appear to affect this assessment context. First of all, many of this teacher's current students are characterized by low SES, needs in oral language, and poor impulse control. It is possible that the relative importance of oral
language and behavior may be a function of her class makeup or it may reflect this teacher's decision making style independent of her class makeup. There is a similar class in the sample. Inspection of Roberta's and Scotty's profiles reveals a similar reliance upon oral language data. The second influence is not discussed here as it contains identifying data.

Influences upon assessment expertise. Like Elaine, Scotty reports relying upon work samples for data when she began teaching. She also states she strictly adhered to all parameters for administering, scoring and interpreting the CRT basal test as a new teacher. She further reports, however, that as she has grown and developed expertise as an assessor, she relies more upon group interactive data and IRI's. "At the beginning I put more weight in the reading level tests than I do now, and I didn't understand how valuable a tool the IRI could be." Scotty does not feel that her reliance upon oral language occurred as a trend in her development. She confidently states that she has always attended to student talk for various cues:

It [oral language] has been something that I have always felt was important because I've always felt that environment and experiences carry a great deal of weight as to whether a child is going to be able to get certain concepts, comprehend the discussions going on in class. . .what they're reading. . .if they haven't ever experienced something or seen it or fully comprehend what it is. . .when they read about it, they won't be able to understand the story. . .so I've always valued experiences and their own language. . .if they can't. . .if their speaking vocabulary is so limited. . .it makes the reading limited too. . .they might even be able to
say the word. . .read the word. . .but I don't think they can comprehend it as easily. It will take them longer. It will have to become part of their speaking vocabulary to a certain extent.

Scotty reminisces about her professional growth and shares an important lesson she has learned through the years: a given set of data may be attributed to several causes. Therefore, it is not possible to say, "I've seen this before; my response is ___." When asked whether she looks for similarities with past experiences, she states:

. . .I don't put a tremendous amount of weight to it. . .because I feel that each child is so unique and has their own reason for it. . .if they can't write a sentence. . .they can only draw a picture. . .sometimes it's emotional. . .sometimes it's academic. . .there are so many reasons. . .I don't put a lot of weight into what has happened in the past.

This is far different from Kitty who states that "if you didn't look for patterns, every year would be brand new and you'd be starting from scratch."

Finally, like other subjects this teacher learned from peers. She is quite specific. Scotty listened to peers value or discount test data in light of what they knew about their students:

. . .I guess hearing other teachers talking about what they thought when they got their test results back. . .if Johnny had made whatever on the test. . .if they knew he could do better. . .or if they. . .just listening to them
talk. . .I learned that you don't necessarily go with a
test score. . .that you judge from their abilities. .
what they're able to do in a group orally. . .

Again, mentorship programs are favored by this teacher. When asked how we can help new teachers to learn how to get a handle on who's learning and who is not, she states:

. . .I guess just always being willing to answer
questions or to listen. . .just to listen. . .if that
teacher has a child that they're talking about. .
.because there is so much talk. . .in general about your
class. I think the mentorship is starting. . .we do that
anyway. . .I think every new teacher there is. . .an
experienced teacher takes that teacher under their wing.
I think that just naturally happens. . .it happened when
I came in. . .I've seen it happen with other new teachers
in our building all along. . .

In addition to formal and informal mentoring of new teachers, Scotty feels the influence of the reading specialist is like a mentor:

. . .if I had a question. . .or the reading specialist
would help me with that. . .but just. . .just to listen.
. . .I would just try things out. . .what do you think
about this. . .should I try this?. . .should I not?. .
have you ever done this?. . .

Recommendations. Scotty provides a comprehensive list of what she feels should be included in a portfolio:

Well, I would want to see writing samples. . .perhaps a
writing sample from the beginning of the school year, the
middle of the year and the end of the year to see what
progress they made. . .at least three writing samples
spread out over the school year. . .I would want to see
different types of writing like creative writing or
applying vocabulary. . .some type of comprehension. .
perhaps where they read a story and answer questions
about what they read. . .a short passage and had to
answer some questions. . .

When reminded that she has mentioned oral reading many times
throughout the interviews and asked about the portfolio in terms of
this concern, she states:

. . .it could be a checksheet as to what types of
mistakes the child made. . .if they're only using the
beginning sound say. . .if they're able to use context
clues. . .what their fluency is. . .if they're fluent
readers. . .if they're relying too much on phonics. .
how long a passage they can read and comprehend what
they have read. . .if they can only do it with a
paragraph. . .if they can do it with an entire story. . .

She does, however, express a concern over the lack of consensus
within the school division on clearly defining learning targets,
selecting instructional materials and methodology, and developing
assessment materials and procedures. She summarizes her concerns and
states:

. . .I guess my concern is that right now there is so
little structure that the gist of what I've gotten is
that each teacher is going to approach this in their own style and manner and way and I'm hearing most teachers talking about novels, chapter books. . .and that's really not the way I want to approach whole language for myself and my classroom. . .and I'm not sure what these teachers mean when they say they're going to use novels. . .if they fully understand they're going to be responsible for incorporating all the skills that in the past have been really mapped out for us. . .and that we're going to have to somehow incorporate these skills into our. . .reading of this chapter book or novel. . .

This statement appears to place Scotty on the change continuum described by Borko, Flory and Cumbo (1993), as an assimilator. That is, Scotty appears to have assimilated some of the whole language methodology into her existing instructional program. However, the underlying belief system providing focus is one of literacy as an accumulation of mastery skills. It does not appear, therefore, that she has confronted her basic belief system regarding how children acquire literacy. Therefore, in order for Scotty to make more than superficial methodological changes, ongoing inservice and support are needed that deal with the underlying philosophy of teaching and learning.

Scotty states that without some standardization she fears her assessments may not be valued by others:

Right now I think it would be more difficult because I think each teacher is going to have to come up with their own way to assessing. . .and as I mentioned, I would very much like to know that a child can read a passage and
comprehend. . . so my feel right now is that I'm going to have to come up with the passage and the questions. . .

which, of course, is more work for the teacher. . . and I don't even know that next year's teacher will put any weight into that. . . so it will be important to me but not perhaps to that person.

**Work samples.** Work samples provided include L.'s generated sentences with vocabulary words. The teacher comments on the student's oral language and his very short sentences. L.'s CRT basal test follows. The teacher questions the validity of the score (75%) and states he may have guessed. However, after closer inspection of the content, she declares the readability on this sample is "easier," and therefore, he scored higher. Finally, activity pages for Snow White complete this packet. This is in sharp contrast to the work packet for C. who is in the top group. There are 23 pages. All are either workbook pages or dittoed sheets. The majority of these are perfect papers.

**Contexts for the think aloud and classroom observation.** The setting chosen for the think aloud was a nonsense sentence activity. This researcher was invited to observe a testing session. Both of these activities have been previously discussed.

**Stacy's Assessment Profile**

**Introduction.** Stacy's school population could be characterized as high SES and a high degree of parent involvement. Her classroom was observed to be quiet, teacher directed, and structured. Children were occupied with paper-pencil tasks during small group instruction. The instructional group sat on the carpet near the teacher. A creative dramatics activity utilized overheads projecting scenery
onto a sheet hung in the front of the class. Some student work was
displayed. The classroom was clean and orderly.

Table 7 records the frequency of data sources stated in all
interviews for this subject. The total number of data sources
identified was 275. A percentage of total responses is reported for
each data category.
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Recognition</td>
<td>49</td>
<td>18%</td>
</tr>
<tr>
<td>Observation</td>
<td>25</td>
<td>9%</td>
</tr>
<tr>
<td>Rate of Learning/Need for Repetition</td>
<td>24</td>
<td>9%</td>
</tr>
<tr>
<td>Comparison of Data</td>
<td>22</td>
<td>8%</td>
</tr>
<tr>
<td>Comprehension</td>
<td>21</td>
<td>8%</td>
</tr>
<tr>
<td>Oral Responses</td>
<td>13</td>
<td>5%</td>
</tr>
<tr>
<td>Behavior/Work Habits</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Retention of Skills</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Grade Level as Reference</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Home Background</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Basal Test Data</td>
<td>10</td>
<td>4%</td>
</tr>
<tr>
<td>Ability</td>
<td>8</td>
<td>3%</td>
</tr>
<tr>
<td>Prior Educational Data</td>
<td>8</td>
<td>3%</td>
</tr>
<tr>
<td>Oral Language</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td>Written Language</td>
<td>6</td>
<td>2%</td>
</tr>
<tr>
<td>Data from Reading Specialist</td>
<td>5</td>
<td>2%</td>
</tr>
<tr>
<td>Processing Skills</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Work Samples</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Testing Modifications</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Peer Coaching Data</td>
<td>3</td>
<td>1%</td>
</tr>
</tbody>
</table>

Table continues
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intuition</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>Requests for Assistance</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Status as Retainee</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Motivation</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Standardized Test Data</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Articulation Disorder</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Willingness to Read</td>
<td>1</td>
<td>*</td>
</tr>
<tr>
<td>Total number of data sources</td>
<td>275</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** * = Less than 1%.
Highest frequency data. This subject made a total of 275 responses. Of these 49 or 18% concern word recognition. This is the second highest percentage of reliance upon this data source. However, of greater importance is the nature of her comments in this area. Approximately one-third of her statements reveal some attempt at error analysis. For example, when a child does not produce any response to the word what, she classifies this differently from other miscues. She explains:

We went over words for like. . .and one of the words for this week was what. . .and he read the word today and he didn't know what the word was. . .and I said, "R., this is one of our spelling words". . .and he still couldn't tell me what the word was. . .you know. . .and we do all kinds of things with the spelling words all week and he still couldn't tell me what the word was. . .

This suggests that she classifies the types of errors students make in terms of amount of instructional time spent on given vocabulary. She displays this same reasoning with J.'s errors:

We have this one little boy. . .J. . .and he's been brought to the team. . .and he's still. . .he just can't remember. . .and hold it in his head. . .like the beginning words. . .like what and where. . .and which and why. . .and we keep going over them and going over them and it's like they just won't stay. . .

What is also of interest here is that she does not comment on the similarity of these words as a category of error. A more classic example of error analysis, however, is contained in Stacy's
discussion of J.'s strengths:

I think he's definitely more auditory than visual. . .um.
. . that just from his reading ability. . .you know I
don't know if the letters are being mixed up in his head.
. .you know, when he goes to read a word. . .um. .
.'cause he'll read a word that doesn't even look like
what he's saying a lot of times. . .

This suggests that she compares the child's miscue to the word
for information about the strategies J. is using to decode. It is
interesting that while J. and R. display similar decoding
proficiency, Stacy suspects J. has more innate ability due to his
relative strength in math, his intense interest in science, and an
error pattern that suggests letter sequencing and reversal
difficulties. She explains further:

R. is more of a. . .just a flat profile. . .I mean. .
.he's below in like reading. . .math. . .every area. .
.whereas J. . .he's weak in math too. . .but the
differences. . .I guess the variances are greater in his
reading. . .because I just see more difficulty in reading
than I do in math. . .um. . .so I think his area of
weakness is going to be decoding and reading
comprehension. . .whereas his math ability I think is
there. . .well, you know he's always trying to answer the
questions in science and social studies. . .and not that
he gets all of them right all the time. . .but he's more
interested in that. . .whereas reading. . .he just finds
it really difficult. . .and with science and social
studies we don't do a lot of reading. . .it's more a
whole group type activity. . .and he's more involved with that. . .and math. . .he picks up on that. . .you know it's going in. . .but it's just not. . .going in the right spot. . .and coming out wrong. . .but I would think there's something that is probably not processing right for J. . .

A finer grained analysis of Stacy's comments in the area of word recognition reveal some unique profile characteristics. Nine text segments state that she values the rate with which students acquire sight vocabulary, as well as their short and long term retention of vocabulary. The assessment behavior that supports this aspect of her decision making style is documentation of errors both in isolation and in context. She states:

Some days he will know his words. . .some days he's just totally forgotten. . .the same words he knew a couple of days before. . .I usually keep a list of the words that they miss and then we keep going back over those words that he's missed. . .you know. . .to see when he's learned these words. . .

When she is speaking of another student, she states:

. . .he reads a lot more fluently than some of the other ones. . .and if he doesn't know a word. . .by the next day. . .he'll keep going over them until he knows them. . .because he doesn't like having a list of words. . .that he doesn't know. . .he'll keep working on them. . .

Thus, it appears that she makes many written records that she
can reflect on later regarding error patterns, phonemes missed, etc. She uses error analysis strategies with students' invented spelling in a similar fashion. For example, J. spelled very as f-e-r-y. Stacy comments:

it could be. . .I'm not sure about that one. . .I think it could be something to do with his auditory perception of the word, you know. . .because it's very not ferv. . .

Stacy's documentation of word recognition errors, formal and informal, is kept in a student's reading folder. She emphatically responds that she would share this information with parents:

I would show them the reading test. . .and the word recognition tests that were given to him. . .and the stories. . .probably just those. . .because the papers go home weekly. . .so they get to see those. . .

Stacy is also unique in her specificity regarding oral reading features. She was observed during a creative dramatics activity. Students used a script and they had read the material several times previously. Few word identification errors were detected. Stacy, however, encouraged students to use appropriate intonation, expression, and phrasing. For example, while encouraging expression, she said to students, "Now if you had a handful of jewels, how would you say it?" When questioned by the researcher about why she chose this activity for observation, she responded, "well, it usually will show their fluency. . .and their rate of reading. . .it will show me a lot of words that they keep missing, you know. . .time after time."

Finally, Stacy includes oral reading in her September sizing-up activities. This is consistent among subjects. She explains her strategies:
Always at the beginning of the year, we'll read with them and go over vocabulary words. and kinda see how their fluency is in reading. but I read with them to kinda see. go over vocabulary from previous units they've already been on to see how much retention is there.

Comparison of data. Stacy's self-reported style is to use multiple data sources: decoding, oral responses, tests, and creative writing. She uses all of these to validate attributions she makes about students during the interactive classroom context. Her analysis of data, she states again and again, is to determine rate of new learning and retention of knowledge. This is consistent with frequencies of data sources with several important exceptions.

When asked whether she values oral responses in the instructional group, or independent work samples obtained after instruction, she chooses the latter and explains:

usually something that they've done on their own tells me more. like if I would have given them a story and do comprehension on their own. or if they've read the story and just ask everybody the same questions, they would have to write the answers down rather than just ask anybody and they would raise their hand. and call on them on the spot. that way everybody had the same question. rather than. with 13 in the group. they all know that only one or two are going to get called on. the chances of them being called on are very slim.
This appears to suggest that she is trying to accumulate a broader sample of data on their ability to answer comprehension questions than a single question per instructional period would allow. Therefore, she sometimes has all students write the answer to the comprehension questions while they are in the instructional group. She uses these data for later reflection.

However, when one considers the actual frequencies of data counted, only four out of 275 responses deal with work samples. On the other hand, over half of her data sources are derived from interactive teaching. It is not clear whether she failed to report this strategy of writing out comprehension responses while in instructional group to be included in work samples, or whether she considers these data as interactive.

It appears that the data in question are collected intermittently while the instructional group proceeds. Therefore, students are assisted in constructing meaning from the text. Their prior knowledge is activated by discussion. Vocabulary is taught and reinforced. Indeed, prior concepts essential to understanding the content of the story are presented. It is not clear whether these data are truly an independent work sample or merely students' recorded responses within an instructional sequence. What is known is that Stacy uses these data in a formative manner: reteaching is focused upon diagnosed needs.

Behavioral data. Stacy's behavioral comments appear to be characterized by a unique focus and specificity. Of the 20 comments regarding work habits or behavior, approximately half were positive, usually noting a parallel improvement in affect or motivation and a similar change in achievement. In one example she states:

...His motivation factor is real high. . he just. .

.he really wants to please. .you, he'll do anything for
me. . .he's always asking, "What can I do for you? What can I do for you?". . . and just real willing to please. . .and um. . .I think our counselor had talked to the counselor at his previous school and she was telling me about how he has changed and how he's hugging everybody and she said that he never did that. . .I mean never!. . .I mean the whole time he was at the other school. . .he would never hug anybody. . .and he just. . .like I said. . .he nearly tackles me and knocks me on the ground. . .he just comes up and hugs me. . .he has just come so far. . .

Stacy states that she monitors during instructional groups for inattention and for those students engaging in self-stimulating behaviors such as rotating pencils and playing with shoelaces. Her response to student distractibility is to provide structure and predictability within the instructional context. She is aware of the inability of some students to screen out distractions as she states:

. . .This is the kind of group I have to remind to keep on track. . .they have a tendency to be real. . .they let outside noises bother them. . .they're distracted real easily. . .and any cue to just keep them. . .we go through the same routine over and over again. . .but when they're reading. . .they have a hard time focusing. . .they just have a hard time. . .focusing in on what they're supposed to be doing. . .and you know, they start fiddling with things. . .and their shoelaces and K. was playing with his hands. . .and it's non-stop. . .
Thus, it appears that Stacy's behavioral comments are tied to learning and reflect strategies to minimize disruption and improve focus. There are several possible explanations for these unique characteristics within this profile. The video tape and classroom observation for this teacher reveal very little attention diverted to behavior management during the classroom sequences recorded. Additional observations in this classroom on other occasions reveal that this pattern is representative. Stacy appears to devote a minimal amount of instructional time to managing behavior. Student behavior during all observations, however, could be characterized as consistently on task. Thus, it appears that this teacher uses highly structured and successful classroom management strategies that significantly reduce instructional time diverted to discipline. This may be attributed to her professional training and experience. These are not discussed here as they contain identifiable data.

**Data collection, class makeup, and methodology.** Stacy states a relationship between class makeup and methodology. When asked where she finds herself on the continuum between whole language and basals she responds, "I guess for this year it's probably more at the other end toward the basals...just because of the students that I have." In contrast, it is interesting to note that Scotty identified herself as utilizing more whole language and language experience methodology because of her students' needs. A closer look reveals that these two groups of students display similar needs. They are characterized by limited oral language, severe decoding needs, and limited prior knowledge. Yet, one teacher responded to these needs with basal methodology and the other with integrated language methodology.

However, Stacy's response is more representative of the subjects in this study. When presented with students with severe decoding difficulties, they express a tendency to rely upon the basal's sequence of decoding skills to insure they have "done
everything they can" for these needy students. In explaining why she must not deviate too far from the basal Stacy declares:

...I wanted to do a lot more with whole language but because of the kids and their weaknesses I just felt like they needed to get a handle on reading and felt they had a hard time decoding words. ...um...you know, comprehending stories...

It should be emphasized that Stacy's data collection for these students is qualitatively different from students in her class perceived to be on or above grade level and proficient at word identification tasks. There is no evidence she uses flash cards, word lists or retesting of vocabulary previously learned with other students. Nor does she attend to the rate with which they learn or retain vocabulary. Thus, class makeup appears to affect Stacy's data collection.

**Data collection and external mandates.** In another interview, however, Stacy provides a different rationale for her orientation toward basals:

...Since we're still responsible for giving the HBJ unit tests...I felt like if I didn't spend a lot of time on the skills, there's no way they would pass the tests...a lot of them...don't pass the tests and we do spend a lot of time...on the skills...so...like I said, I'm geared more to the high end toward the basals just because of my students this year...and only because we're still responsible for giving the unit test...if we weren't responsible for that then I'd probably be geared more toward whole language...but I felt since
we're still responsible I'd better do more toward the basal.

**Observational data.** Stacy expresses confidence in her intuitive judgments regarding her students' learning based upon observation. "That's why I think seeing them day after day, you can really say, 'yeah, I think they really know the *ax* sound or they don't know it." This is consistent with her relatively high percentage of reliance upon observation.

**Comprehension data.** Stacy's reliance upon comprehension as a data source is second only to Scotty's. While her word recognition comments focus largely on remedial students, her comprehension comments were directed toward explaining how she knows which students are "getting more" from a story. This getting-more behavior appears to include the ability to recall without rereading the text. She recalls an example: "She's able to answer a lot of the questions when a lot of the other students can't." In addition, she describes "seeing more" or understanding the author's message and purpose. She further explains:

> ...She'll go back and look and she sees a lot more into the story. If there's some hidden message in the story, she usually can see it...

Finally, she values the ability to understand humor and states this is often difficult for the below grade level child. Her comments are similar to several other subjects:

> ...If there's a pun or a joke...a lot of the kids who are below grade level don't usually see those little puns and jokes and the funny things in the stories...I know what it was...there was this one story we read...
about this little girl who liked to do gymnastics. .
and her father wanted to teach her how to play tennis. .
and at the end he said, "well, maybe someday when you
get older I'll teach you how to play tennis. . .and it
can be your second game," and she turns around and says,
"yeah, I'll be glad to teach you gymnastics for your
second sport." And she was the only one that got that. .
nobody else in the whole group did. .

Thus, it appears that while Stacy does appear to respond globally
when talking about comprehension, her examples are specific.

Stacy uses comprehension data in her September sizing-up.
activities. "I read with them to kinda see. . .and I ask them a few
questions to kinda see, you know, where they're coming from as far as
comprehension." She also demonstrates her understanding of the
relation between reading rate, phrasing, and comprehension. In
commenting on R.'s needs she says:

. . .His rate is a little choppy and he has a hard time
comprehending and I think a lot of it is because it is so
choppy that it's kinda hard to get the flow. . .you know.
. .of the story a lot of times. . .when they read real
slow. . .he may read a few words that he knows in a
group. . .and then it's a long pause. . .so, I think as a
whole. . .they have a hard time in their understanding of
the story because of that. . .

Oral language data. The frequencies of data sources reveal
that Stacy makes fewer comments about oral language than all other
subjects. She also does not indicate using language as an estimate
of ability as many other subjects do. However, it is interesting
that she chooses a vocabulary activity to be videotaped for the think aloud.

Her comments about vocabulary or word meaning study are typical of other teachers' responses. She addresses students' tendency to define a word using that word. Then she relates oral language to students' background of experiences. She appears to understand students' difficulties in pulling information together to form a concept such as the word *family*. She explains:

...They have a real hard time with giving meanings to words...and being able to put it into a sentence. ...
.you know...that will make the word meaning understandable. Like we were talking about a baby...

From the video tape, one hears the teacher prompting a student after he has provided a partial oral response. On the videotape Stacy says, "...so a baby is a small what?" The child's response is not audible but one hears Stacy apparently elaborate upon his response: "...O.K. You have to take care of it and it can't walk. ..so a young child that's usually..." To the researcher, the teacher states:

...They had a hard time with the word *family*. We went over the vocabulary the day before...and they had a hard time and I said, "well, what is a family?"...and they said, "well, it's people"...and I said, "O.K., it's people and is our class kinda like a family?"...and they said, "well, yeah"...then I said, "tell me some more things about a family"...then they said, "well, the people love each other"...but they had a hard time grasping that it was mom, dad, brother, and
sister. . .so finally, someone said, "it's a brother and a sister". . .and I said, "well, that can be part of a family". . .and it was like I think they have a lot of areas that they just don't have a lot of experience being able to verbalize what words mean. . .not a lot of experiences out in the world. . .to be able to gather all this information to put it together. . .

Thus, it appears that this teacher utilizes response elaboration and provides scaffolding to assist students in making these language connections. The student who gave a complete and adequate response during the vocabulary activity defined signals as "signs that tell us where to go." The teacher remarked that he has more vocabulary than his performance would indicate. Although she notes the difference, she does not extend this to an inference about his innate ability.

Finally, Stacy attends to tone, length of utterance, and grammar or oral and written language. Again, in discussing R., she states:

. . .he doesn't respond orally real well. . .he's real quiet and you have to get him to speak up when he talks because you can barely hear him. . .um. . .he doesn't speak in complete sentences. . .or correct grammar. . .and I attribute a lot of this to his home life. . .because I don't think they speak in correct English either. . .

It is not clear whether grammatical features here result from differences that could be attributed to dialect, ethnicity, native
language, social class, or other factors.

Prior educational data. Stacy and Roberta make 11 and 13 comments about home background respectively. In addition their comments are focused around themes of family stability and its effect on learning. In describing a child's lack of progress, Stacy attributes part of K.'s difficulties to his home situation. In addition, she speaks to the issue of the lack of support within the home setting for reinforcing learning:

... His home life is real bad. ...his mother's not in the family and his grandmother takes care of him. ...and she's trying to get custody. ...you know, from the father. ...because the father's hardly ever around. ...
...and they come to school unfed and not clean. ...and haven't been doing their homework because he was supposed to be taking care of them. ...and he wasn't. ...

These comments relating home background and learning are also only made as an attribution for the poorer readers. Betty, on the other hand, appears to provide home background comments in explaining a judgment independent of achievement. As most subjects, Stacy consults previous report cards. Although she admits doing this at the beginning of the year, she is careful to issue a caveat of possible expectancy bias, and states she forms her judgments independently of these data:

...I usually look at the report cards at the beginning of the year. ...but I'm the type of person. ...I guess I like to form my own judgments of kids. ...and not really go by exactly what it says in here. ...and expect him to do the same thing. I would rather see for myself. ...so
I kinda glance at them. . .and then I'll usually look at them later. . .just to kinda see if it's meeting my expectations and what I think I see. . .in this child.

Other prior educational data used by this teacher include information contained in a student's reading folder. The front of the folder provides a list of texts taught the previous year and dates for each. Stacy states that she scans this for clues about pacing or rate of movement through instructional materials the previous year. She explains what these dates tell her:

...Well, the front tells me that he had really slow progress in first grade. . .and he really didn't cover a whole lot. . .which he should have. . .in first grade. . .so he had a lot of difficulty. . .

For another student, she states:

...his reading folder. . .he progressed exactly where he should have been. . .last year. . .he ended first grade on grade level. . .on exactly where he was supposed to. . .he was just. . .you know, an average kid. . .

This last quotation is even more meaningful when one considers this particular student was found eligible for special education under Public Law 101-476. One might interpret the teacher's comments here as an effort to emphasize this child's progress despite his disability.

Test data. Stacy provides an interesting rationale that sheds additional information on why her profile of data sources only contains four percent of counted responses from basal test data as contrasted with nine percent reliance upon her own observations. She
thinks aloud about test data and intervening variables:

...With the tests like HBJ...they could do really
badly or they could do really well...but that doesn't
necessarily mean they know the skill...they may not
test good...a lot of different complications...so
just going back and giving informal testing...I guess
giving the HBJ tests...I've come to realize that they
can pass the test but that doesn't always mean that
they've mastered the skills...They could be good test
takers or they could have guessed real well...I guess I
have a hard time accepting that because they passed the
test, they really know the skills...that's why I think
seeing them day after day, you can really say...

Thus, it appears that Stacy uses basal test data fully
cognizant of its limitations. She compensates for these by
collecting multiple data sources and inspecting students' data
profiles for discrepant pieces of data. These she verifies with
informal observations and performance assessments. Moreover, she
does not appear troubled by the flaws in this data source. One must
recall that she has previously pointed out reliability threats in
standardized tests as well as work samples. While performance
assessments may have extremely high face validity, she is concerned
about the lack of coverage possible.

As many other subjects, Stacy values narrative comments on
students' folders. She attends to statements about testing
modifications and if a student was retested. Finally, she indicates
that she attends to the comprehension score on the basal test more
than all other data. She talks about what the comprehension score
tells her:
I guess the comprehension part of the test is better than the other parts of the test. because they do have a new story that they're reading and they're not discussing it with the teacher. and so I think that it will show whether they've mastered main idea or whatever.

This subject reports looking for patterns in prior data and even attends to performance as a function of the time of year. Time of year is a unique concern to this teacher. She states:

Yes you want to look at the previous report card to see if they've having trouble with the same things as they've having trouble with you. but you want to see what their pattern is over the whole year.

Stacy's rendering of cumulative folder data is the most thorough of all subjects. She appears to have considered carefully which information she will use and is able to articulate her rationale. For example, she discounts the validity and reliability of the COGAT because it is not a reading test, because attention can depress scores on a listening instrument, and because stability of standardized measures on six year olds is a concern:

As far as test scores. like with the COGAT testing. I mean I look at it. but I guess I don't put a whole lot of emphasis upon the test scores because they can guess and get them right and that sort of thing, so I'd rather see more one on one type evaluation where you can reliably see. and I can see how he performs and with those kinds of tests you just read it to them.
and they're supposed to answer... and hopefully be following along... and so I don't... if they don't feel well or whatever it is... not pay attention... it's a long time for them to be listening...

Records from another school are also discounted as a valid source of information because rubrics and evaluation criteria vary. She explains her reluctance to use these data:

...It's hard to look at a child's cumulative records from another school... just because you don't know their method of evaluation... 'cause some schools use 0's, and you know, VG's and you're not sure what that criteria is based on...

In a truly unique segment of this study, Stacy demonstrates her ability to read and interpret a standardized test profile and then compares that with the student's classroom performance. No other subject demonstrates this skill. In looking at an average student, she states:

...he's just an average second grader who picks up things appropriately and you don't have to spend a whole lot of time remediating... with his COGAT, he scores within the average for verbal and quantitative... but nonverbal was the highest... um... as far as his report card... from first grade... he didn't get all VG's... but he is an average second grader as far as academically... which is what he is doing in second grade... getting B's... with his academics...
In another interview she profiles a student who is below grade level:

... She wasn't here for ITBS last year... here's her test this year. The reading specialist gave it to her. ... the only area that she was even in the average range was math problem solving. ... but all the rest of them were below the national average. ... so she's just. ... but she didn't have any other test scores that came with her. ... 'cause this was just taken this year. ... except for the Metropolitan. ... which you know. ... is in Kindergarten. ... 35th percentile. ... so, anyway, she's just one of those kids. ... flat profile. ... well, we've talked about her in Student Assistance meetings. ... but I know she's a flat profile. ... and with that they've not going to identify her with any other handicap other than speech and that's it. ... so she's one of those kids that you know. ... you have to spend a lot of time with. ... and you hope one day. ... it will click. ... eventually. ...

The concept of flat profile is not unique to this interview. Other subjects in the study use this term to describe students who appear to have global weaknesses in all subject areas. However, this teacher draws data from standardized tests to confirm her judgments based upon daily performance. Other informants appear to base their flat profile attribution upon comparison of performance and upon an ability estimate they derive based upon oral language data. It appears that Stacy's expertise in this area is a result of her training and experience.

Work samples. As noted, Stacy states that she values
independent work samples. However, she provides additional parameters for these samples in order to yield useful data. First of all, she desires a reasonable sample of the behavior. She states a lack of confidence in her ability to generalize from one comprehension response from a student in an instructional group. Secondly, she insures that no peer assistance is provided in order that a reliable performance measure is obtained. She adds:

...like with comprehension like a story...with questions...if they did it...if I made sure that they did it without being able to copy...if I separated them and told them, "this is not something you can ask your neighbor if you're not sure of something"...kinda like when we do our tests...

Finally, she appears to understand the inverse relationship between teacher assistance provided and data that may be used for decision making. For example, she monitors the number of requests for assistance she receives during work that she perceives will be independent for students. In talking about a child who is doing well but is not as independent as she would expect, she states:

...and she still has time when she shows me that she's still really needy as far as not understanding something. 
.basically I get that from...if she has a paper to do...and usually I'll help her with the directions...
.but usually those kids can read the directions and figure out what to do...and a lot of times she'll come up and say, "well, I don't understand what to do"...

Grades. Stacy has also given careful consideration to what she
will grade. Grading appears to be a function of what is tested on the CRT basal test. Typical of most respondents, she states that she does not grade initial teaching or guided practice. She explains her procedures:

...Well, the way that I decide what goes in the grade book is, I check the skills that are going to be tested in that particular unit...and those are the skills that I know have been gone over many many times...and so something I know has been gone over and is not just brand new...I'll usually take grades on those...something that's just been introduced...you know, or is not going to be tested because it's just been introduced...or just been reviewed...maybe one or two times...then I don't take grades...

Thus, it appears that the basal test drives her data collection in this area.

In addition to grading workbook pages, she also gives students comprehension questions in the instructional group. They then write the answers. These are apparently given orally. One infers that she is able to monitor to insure reliability of response. This is a unique example of comprehension assessment and she explains:

...I get grades from lots of different places...like the study books or something I would just give them in the group...when we do a whole group thing...and they have paper and pencil and they're answering questions or doing some stuff together with me...I haven't done a whole lot of board work this year...
When asked whether she can predict how students are going to do on a test, she describes her method of formative assessment:

...Yes, I can predict from the grades. ...I'll know who's still having trouble in a certain area. ...and we can keep going over it and over it. ...and I can pull them aside. ...but they'll still have difficulty on the test. ...you know, with the same thing. ...

Stacy then cautions that grades are not truly reliable as a formative measure because they are mediated by knowledge as well as work habits. In reflecting on K.'s grades, she says:

...he goes from A's to D's to F's. ...but mostly he goes A's, B's and C's. ...and I think on his last report card he ended up with a B average. ...so he's really come a ways. ...the F in the grade book here was on a ABC order thing which I think. ...the next one he did he got an A on. ...and then a D. ... he has a tendency to rush through his work a lot of times. ...and he just wants to get it done. ...and I'll have to give him things to go back over and when I do, he gets it right. ...you can't go by grades totally you have to know the kid. ...but with him. ...I know he rushes. ...so it's like, 'O.K., go back and do it. ...I know you know how to do it'. ...

Thus, it appears that Stacy feels observation is required in order to interpret accurately grades on paper-pencil activities.

Written language data. Writing samples or creative writing are valued by this teacher as a rich data source. She perceives fewer threats to reliability than with other work samples as she states:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
I can probably tell more just in their writing samples. ... than their seatwork. ... only because. ... a lot of time with their seatwork. ... because they’re working in groups at tables a lot of times. ... they tend to look off of other people’s papers. ... and kinda do things more cooperatively. ... which is fine. ... so that’s why I feel like the writing samples are more indicative of what they really can do. ...

Stacy states that she would show these to parents as a representation of the student’s ability to internalize and integrate different parts of knowledge to construct meaning. She sums up her thinking:

Writing is a better measure of what they can really do because there they’re pulling in all their knowledge for all the different areas. ... that they have been surrounded with in school. ... and so that kinda shows me. ... like if they have learned their capitalization and punctuation. ... and grammar and all the different. ... you know, comprehension. ... as far as keeping to the main topic. ... so I think their writing samples would be more of an indication to me. ... to be able to show their parents rather than just their individual seatwork. ...

This researcher acknowledges that this is the second time this subject has referred to the process whereby students integrate various bits of knowledge to generate meaning in the classroom. The previous reference concerned oral language. In addition to examining
creative writing for evidence of skill development, Stacy also attends to grammar. She looks at K.'s rough draft and remarks:

...he still has a lot of grammatical errors...these are stories that they would do on their own...you can see the grammatical errors he still makes. K.'s story reads:..."I had a nice Christmas time. My sister [sic], my brother and I had too many toys. A friend of my [sic] and me was playing with our toys. And we was running to must [sic]."

There are numerous erasures and some errors have already been corrected in the first two sentences. Stacy makes other judgments from students' creative writing. She feels sentence linkage and staying on a topic relate to comprehension. When asked how functional writing provides information about a student's reading development she responds:

...well, being able to put the sentences together...tells me that his comprehension is a lot better...and he's also keeping to one subject...a lot of times they like to flip flop back and forth...you know, if I give them a topic about winter...you know, they may start talking about something they did in the summer...and he's able to keep on the subject which is relating to main idea and giving me detail sentences...

She also believes that students who engage in wide recreational reading and are fluent readers, produce creative writing that is characterized by better word choices and varied sentence patterns. In a very specific example, Stacy predicts the difference in written
language she would expect from a grade level and then an above grade level reader:

...Well, writing tells me about word knowledge too. ...because they're able to pull in a lot of vocabulary which would be words they had seen in their stories and so he's bringing, you know, a lot of times they'll bring in a variety of words...than using, "I went to school, I went to, I went to"...rather than everything always being the same...and not just the same word over and over...they would say, "Wow! We had a great day at school!"

Finally, her analysis of J.'s story reveals that Stacy is considering multiple sources of data at one time: ability to write a complete sentence, spelling errors, error analysis of spelling errors, comparison of spelling errors with previous spelling words taught, classification of spelling error pattern, variety of sentence patterns, sequential development of topic, and maturity of the student. J.'s first writing sample follows with errors intact:

I got a niaf. My Brohr got a fotBoll. I had fon. I got to see the crems keros. I got a jolp sat. I got the shrnt fo The CHICago Bulls." A translation follows: "I got a knife. My brother got a football. I had fun. I got to see the Christmas carols. I got a jump suit. I got the shirt of the Chicago Bulls.

When asked what she concludes, Stacy responds:

...Well, still a lot of spelling errors...well, what it indicates to me...he can put the words into
sentences. . .you know. . .to make them make sense. . .um. . .but he has a lot of trouble with vowel sounds as far as being able to distinguish them. . .um. . .and being able to write them down for his spelling words. . .like with jump. . .we have the initial and the final consonants but the vowel sounds are real difficult. . .like knife or fun. . .football is not too bad. . .some of these words are words he's had in spelling. . .and they'll just learn them for the time and then it's just gone. . .well, he kept with the subject which is good, you know. . .as far as main idea. . .there's just a lot of difficulty with him. . .you know. . .repetition of the same kind of sentence. . .which indicates a lot of immaturity and. . .not. . .um. . .how to explain it. . .just. . .I don't know. . .this repetitive sentences. . .which to me is not a real big indication of growth as far as his comprehension. . .

Thus, Stacy expresses more criteria for looking at written language than all other subjects. Her list includes: sentence patterns, variety of sentence openers, grammar, analysis of spelling errors, punctuation, word choice, development of a topic, and sentence linkage. J.'s second writing sample follows:

In the wentr et is cold. I like wentr. Et is fun in the wentr. In the wentr et is fery cold and freing.

In commenting on his spelling of very, Stacy thinks aloud:

It could be. . .um. . .I'm not sure about that one. . .I think it could be something to do with his auditory
perception of the word, you know. . .because it's v-e-r-y, not f-e-r-y. . .I don't know. . .he baffles me. . .he really does. . .because he'll do real well with one thing and then his reading. . .it's just not there. . .there's ability there but there's just something that's not there. . .

Thus, in considering Stacy's analysis of student writing, it appears that this teacher has carefully considered the aspects of creative writing that parallel reading development. However, for J., who possesses good oral language, concept development, and stronger math skills, she still sees immaturity in his written language. She is at a loss to explain this. She appears to conclude, therefore, that for this child, written language is not a valid reflection of his ability or his oral language. It should also be noted that this teacher stated earlier that she feels J. may have a processing deficit.

Influences upon assessment expertise. When asked to describe her development as an assessor in reading, Stacy readily admits that her early years as a teacher included assessment by experimentation:

. . .I used trial and error. . .as far as trying things out. . .and saying "I'll never do this again". . .or. . ."yeah, I'll try this technique again."

In addition, she reports influences from her peers:

. . .a little bit from the teachers. . .you know. . .talking to them as far as what kind of strategies they use. . .you know, to assess children. . .She also
mentions coursework in assessment. It is noteworthy that she even mentions this influence. Recent surveys of graduate coursework outlines in this area suggest an overreliance on standardized measures resulting in a lack of relevance for practitioners (Stiggins and Conklin, 1992).

Finally, Stacy states that additional influences upon her development as an assessor include the reading specialist and collaboration with peers. She explains her interactions with the reading specialist:

. . .If I had some concerns about some of the kids in the group. . .then I would go and tell the reading specialist, "here are my concerns". . .and tell her some of the things I've done and then get suggestions from her. . .oh, well, a lot of times if I do have concerns about them. . .then she'll take them individually. . .and test them. . .you know, do some assessment with her. . .one on one. . .or if I tell her the assessment stuff that I've done with them. . .and say, "I'm still not sure what to do with them," then she may give me some more suggestions as far as what to do. . .sometimes they may do one thing with somebody else and something else with me. . .so a lot of time . . .I let her see them just to see the effect. . .I've done that with the teacher across the hall. . .sent them over. . .and said, "read with this child and see if you're seeing and hearing the
same things that I'm seeing."

This is a common comment about mentors and collaborators: they need to go into the new teacher's room and look at her children or her data in order for the teacher to gain new insights or skills in interpreting student achievement data. Although Stacy acknowledges that student records reveal patterns of achievement, she denies utilizing patterns for diagnosis. She states firmly:

...I don't look at patterns in kids...sometimes if they're real similar...but a lot of times I find they're all so different...they may all be lumped as far as below grade level...and then...I think I'll see some similarities but I don't...I think they're all so different...

In a pattern similar to other respondents in this study, Stacy's recommendations for inservice include development of mentoring programs for the new teacher:

...with the mentor program...we could do that...having that teacher work with the new teacher...to give her some ideas as far as assessments that she could use...probably show her some assessments that have worked in different circumstances...and she could say, "oh, I've tried this one...you can try it...and it may work for you and it may not"...even going through an assessment with her group and then talking to somebody and saying, "now here is what I found and, I think this
is what it is telling me". . .and have somebody at least listen. . .to what she thinks is wrong or right or whatever. . .and have them say, "yeah, I agree with that. . .or I see this. . .this example shows me that the student doesn't know short vowels or whatever". . .going through it, step by step, with the new teacher. . .I think that's where a mentor program might come in really handy. . .

She also favors formal inservice for new teachers to build a knowledge base:

. . .probably some inservice. . .maybe pulling the first year teachers out. . .showing them different assessments that other teachers have used in the past. . .

Finally, in developing the expertise of the beginning teacher as an assessor, Stacy recognizes the importance of the ability to interpret data as a key component. She demonstrates in her comments that ecological assessments in the classroom context are mediated by a myriad of extraneous variables. The freight that each of these variables carries in accounting for observed performances in reading must be teased out in a systematic manner by the teacher. All of this must occur before parent conference day. Stacy's interviews conclude with this researcher's favorite quotation:

. . .oh yeah, you could assess from here to June 18, and not know what to do with it. . .so yeah. . .you have to know what it's going to show you.
Think aloud and observation context. The setting chosen for the think aloud was a vocabulary activity. The teacher provided response elaboration and scaffolding to elicit more oral responses to create the concept of the words baby and family. The researcher was invited to observe a creative dramatics activity. Children were practicing for a play by reading aloud from scripts. The teacher's comments were focused upon expression, intonation, phrasing, and inflection in order to convey meaning and feeling.

Kitty's Assessment Profile

Introduction. Kitty's classroom was warm and comfortable. Pillows, quilts, and many tradebooks were found around the room. Children engaged in recreational reading. They talked and shared books. The teacher appeared to facilitate rather than direct the classroom, and children moved freely. Student writing, art work, and projects were on display. Social studies and science themes were evident from these displays. Group authored charts were visible.

Table 4.8 records the frequency of data sources stated in all interviews for this subject. The total number of data sources identified was 165. A percentage of total responses is reported for each data category.
Table 8

Data Frequencies for Kitty and Percentages of Responses by Category

<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Recognition</td>
<td>35</td>
<td>21%</td>
</tr>
<tr>
<td>Work Samples</td>
<td>12</td>
<td>7%</td>
</tr>
<tr>
<td>Comparison of Data</td>
<td>11</td>
<td>7%</td>
</tr>
<tr>
<td>Comprehension</td>
<td>10</td>
<td>6%</td>
</tr>
<tr>
<td>Observation</td>
<td>10</td>
<td>6%</td>
</tr>
<tr>
<td>Oral Language</td>
<td>9</td>
<td>6%</td>
</tr>
<tr>
<td>Motivation</td>
<td>9</td>
<td>6%</td>
</tr>
<tr>
<td>Prior Educational Data</td>
<td>9</td>
<td>6%</td>
</tr>
<tr>
<td>Behavior/Work Habits</td>
<td>8</td>
<td>5%</td>
</tr>
<tr>
<td>Oral Responses</td>
<td>7</td>
<td>4%</td>
</tr>
<tr>
<td>Written Language</td>
<td>6</td>
<td>4%</td>
</tr>
<tr>
<td>Basal Test Data</td>
<td>5</td>
<td>4%</td>
</tr>
<tr>
<td>Auditory Processing</td>
<td>4</td>
<td>3%</td>
</tr>
<tr>
<td>Conferences with Students</td>
<td>4</td>
<td>3%</td>
</tr>
<tr>
<td>Rate of Learning</td>
<td>3</td>
<td>2%</td>
</tr>
<tr>
<td>Peer Coaching Data</td>
<td>3</td>
<td>2%</td>
</tr>
<tr>
<td>Data from Reading Specialist</td>
<td>3</td>
<td>2%</td>
</tr>
<tr>
<td>Performance Assessment</td>
<td>3</td>
<td>2%</td>
</tr>
<tr>
<td>Independent Reading Choices</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Standardized Test Data</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>Status as Retainee</td>
<td>2</td>
<td>1%</td>
</tr>
</tbody>
</table>

Table continues
<table>
<thead>
<tr>
<th>Data Category</th>
<th>N</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skills</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Willingness to Read</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Intuition</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Articulation Disorder</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>L.D.</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Ability</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Medication/ADHD</td>
<td>1</td>
<td>1%</td>
</tr>
<tr>
<td>Total number of data sources:</td>
<td>165</td>
<td></td>
</tr>
</tbody>
</table>
The discussion of Kitty's profile focuses first upon those data categories which occur most frequently in the interviews. Kitty's interviews contain 165 references to data sources. This is the fewest number of all respondents and only half of Scotty's 329 responses. Kitty's brevity is not perceived as a function of the interviewer-informant relationship but rather to her own personality traits of shyness and reluctance to take an expert stance on a topic. She makes several statements about her own skills as an assessor that suggest she feels uncomfortable taking the sole responsibility for assessing her students' reading achievement or being accountable to parents. For example, when asked what she does with a great quantity of data she replies, "Cross my fingers and pray!" When asked for ideas about helping young teachers who would be like she was during her first year, she states, "I'd feel sorry for them." Finally, when asked if she looks for patterns in student achievement or data profiles, she responds, "I kinda bumble through life sometimes but I really believe if I didn't look for patterns in how things happen, every class would be new."

Thus, from a cursory inspection of Kitty's responses, she appears to be a somewhat disorganized and impulsive respondent. However, a careful analysis of her responses reveals that she is deliberate in her assessment strategies. The observed differences between this teacher and other respondents appear to be in the area of methodology. This classroom is characterized by more movement of students, a very informal atmosphere, and little evidence of basal methodology. Children engage in reading alone and in small groups or with partners. Journals are used daily. Reading response logs were offered as work samples by this teacher. She uses thematic units and literature to focus instruction. Many creative projects from past units were on display including a quilt and class books. Many
students had assisted in creating a large book character for the bulletin board out of paper, paint, crepe paper, and fabric scraps. Large pillows were scattered around the room and children felt comfortable laying down in a corner to read.

**Word recognition data.** Kitty has the largest percentage of responses regarding word recognition of all respondents. These data are possibly skewed because she chose an instructional activity involving metacognition to be videotaped for the think aloud. Children were presented with flash cards to decode. They were then asked to share what strategies they used to decode the word. However, Stacy also chose an oral reading activity for videotaping and other subjects included oral reading as part of the instructional sequence. Another explanation for this profile characteristic is that Kitty engages in more error analysis discussion than other respondents (with the exception of Stacy). Her thinking aloud concerning student word recognition errors accounts for many of the data frequencies recorded.

Although this teacher appears to use many whole language strategies, the majority of her word recognition comments are made about decoding in isolation:

...and then I also want them to read the words in isolation. ...and I have flash cards for each level and we do flash cards every few days. ...

There are numerous examples of error analysis:

...C. knows the sounds but can't apply them to words. ...there are sounds out there but...that they were making no connection whatsoever between the sounds and the word. ...they can give me the beginning /p/ sound and then give
me a word that begins with /s/...or a /t/...and I could not find any consistency...

Kitty uses documentation strategies for errors including phonetic transcription. She explains, "...when they call the word incorrectly, I would write down what they say...and if I don't do anything then it's correct." The only feature of oral reading mentioned was "fluency." Oral reading is included as a September sizing-up strategy.

Although she mentions observing students to determine which phonic elements they know, she also states that only one student in her below grade level group can use phonics appropriately as a decoding strategy. Her goal does not appear to teach more phonics to these children but to identify those for whom this is an appropriate strategy. Her thinking here seems to suggest that she does plan for these students in light of the needs she has diagnosed, and not solely in content chunks to be covered as reported by Yinger (1980). She explains her thinking:

...K. is the only one that seems to be able to make any sense out of the decoding skills...C. has been tutored in an intensive phonics program but is unable to apply skills learned. He thinks he has a good understanding of what he is supposed to do with sounds...but he doesn't.

Thus, matching methodology to a child's particular learning style or needs appears to be her motivation in this area:

...When I got them, we first just starting working on decoding skills...these children do not have decoding skills...and I decided that they were probably never going to get them if I stood on my head and gargled...
peanut butter!

She extends her interest in what phonic elements students know and can use effectively, to what other strategies they use to decode. The lesson Kitty chose for videotaping for the think aloud consisted of having students decode words in isolation and then explain to the group which strategies they used. Thus, it appears that she used a group assessment activity as an instructional one. From observation in group, she already knows that these students' initial strategy is to ask someone to tell them the word. She reasons aloud:

...what I want them to do is just to be constantly thinking of lots of different ways to approach a word. ...that they don't know and that's why I keep asking them how they do it. ...because this group will come to a word they don't know and would have no clues whatsoever for figuring it out. ...they just stop and that is it. ...so I started asking them how they do what they're doing. ...I've paired them with other children in the room. ...have them show ways to attack the words. ...

She then critiques their strategies to herself:

...There have been in the past some very incorrect strategies. Their number one strategy is ask. ...that was number one. ...that was when we first started doing it. ...asking well, how do you know a word that you don't know. ...then the next thing they would all say is sound it out. ...but the way they were sounding it out. ...was. ...
The unfinished thought here is provided in another quotation where she notices that a student produces the response *ug* for *use*. She states, "see, he was looking for the little word...now he doesn't know what to do with it." This suggests that she feels when her students attempt to use structural analysis or sound words out, their strategies frequently break down.

**Work samples.** Kitty and Amy's profiles indicate they rely more heavily on work samples than other respondents. However, Kitty is unique in stating that she would prefer to design her own worksheets in order to feel more confident in her measurements. She talks about her criteria in designing worksheets:

> . . .if I wanted to show them that . . .I would do a paper with vocabulary and comprehension. . .because those are the two things I'm really looking for. . .I'd rather just have them read a sentence. . .take the vocabulary word and write a sentence. . .on their reading level or a lower reading level. . .that would include the vocabulary word and have them read the sentence and then also give them a list of words in isolation.

There are several important concepts presented in Kitty's text segment above. First of all, it is likely that Kitty would not use the basal worksheets because they would not meet the needs of this below grade level group. Secondly, she is concerned with the readability of independent work and knows it should be below their instructional level for them to be successful. Finally, she states her preference for cloze or fill in the blank activities with a word bank for measuring their knowledge of vocabulary.

In another interview Kitty compares a child's performance on one of these teacher-made vocabulary sheets with her expectations for
that performance. Here she reasons:

...I'm not sure with L. whether he's even applying himself...whether he's really reading the sentences when they're on the board or whether he's sticking words...in there...and not trying to read them...because he can read them and read them correctly...

This appears to indicate that Kitty has designed this boardwork utilizing words she has heard L. read. Therefore, she states with confidence that his lack of performance cannot be due to word recognition difficulties. She then hypothesizes that his performance is due to not reading the sentences and just "sticking words" in the blanks. A teacher who perceives that a student has refused to perform or rushes through work will most likely respond to that student regarding the work in a different manner than the student she hypothesizes had difficulty with the readability or the concepts or both.

Decision making style. Kitty's decision making style is consistent with Ken Goodman's description of a "kid watcher." She prefaxes a description of her data base with the disclaimer:

...I know it sounds silly...but it's just being with them and listening to them and talking with them and watching them work and watching them interact with the other children...and I was listening to them read...observing the types of things they read...

One inconsistency noted here is that she does not mention work samples and yet, there are 12 comments about this data source in her interviews. In addition, she does state clearly that she values interactive data from the group setting more than any independent
paper-pencil data. Other inconsistencies include an absence of statements in her interviews about independent reading choices or peer interaction data. When compared to her stated style above, one wonders if she processes these data automatically and failed to report them during the interview.

Kitty also uses prior educational data to support her decisions in a manner similar to other respondents. In addition, she uses interactive data in the instructional group as formative data and waits to give graded assessments until she perceives the majority of the group has mastered a given concept. In this respect, she is most like Amy, who articulates her reason for waiting is to insure that collection of formal assessment data does not lower students' self-esteem.

**Unique factors affecting this assessment context.** Kitty introduces a variable that affects her decision making that has not been discussed by any other subject. She states that the variance of her class may affect the accuracy of her ability to know who is learning and who is not:

...just by being around them...you can't be in a classroom of children without picking up the two extremes...the ones who are doing really well and those who are not...

Thus, it appears that she feels teachers are more accurate in predicting the achievement of students in the top and bottom ranges of achievement. This is consistent with the findings of Gaines and Davis (1990) who reported teachers could more accurately predict the achievement of students whose standardized scores fell in the top or bottom quartiles. Following this finding to a logical conclusion, is it possible that teachers have greater difficulty assessing the
gradients of learning that occur near the median of the class? Additional research may be needed in this area. Chittenden (personal correspondence, 1993) reports some encouraging findings: as teachers are trained in authentic assessment methods, they become more attentive to subtle differences in student achievement within their classroom.

Data collection and class makeup. This teacher's data collection does appear to be affected by class makeup. Kitty's concern with word recognition seems to be a function of reading level. She only collects word recognition data and performance assessment documentation on her below grade level group for whom she perceives decoding to be a major need. "I have flash cards and we do flash cards every few days with this group. . .I don't do flash cards with my other group at all." Field notes for the first interview indicate that Kitty uses basal materials with the six students who are significantly below grade level. She states that she tried whole class teaching with literature and observed that it was not appropriate for the needs of these students.

Data collection and orientation toward methodology. Although Kitty displays more whole language strategies and uses literature exclusively with the majority of her class, her data collection appears to be driven by her choice of instructional methodology. This is supported by her skill in modifying basal materials, making her own worksheets to accompany basal materials, and differentiating collection of data according to perceived needs. In the past when utilizing a basal exclusively for instruction, this teacher displayed creativity in creating unique extension activities and tying stories with themes. Kitty's personality appears suited for a literature-based classroom; however, her data collection seems to be a reasoned response to perceived student's needs, independent of instructional methodology.
Observational data. Kitty sums up how she values this category of data:

...but what really tells me...what I place the most stock in...is their daily performance in the group...
each day...how they perform...how enthusiastic they are... 

This is consistent with her data profile. Over half of Kitty's identified data sources are derived from interactive instruction in the classroom context. What appears to be unique is her attention to documentation. She states that she records word recognition errors phonetically. She also states that she is collecting miscues for a learning disabled student in order to share these with the specialist.

Comprehension data. This subject's definition of comprehension suggests she is closer to a constructivist view of literacy than other subjects. She wants students to make meaning while reading. She wants them to enjoy reading and adopt it as a tool for life. She judges whether this is happening from students' oral responses in the instructional group:

...I can just tell by his responses...he doesn't get the idea of what we are discussing. He'll give responses like the "pig was little" or "his eyes were blue"...He doesn't grasp the meaning.

Kitty includes other comprehension features in her responses:

...they did not do as well with reasoning, comprehension and vocabulary...not only reading the words but knowing what they meant...
In another interview, she states, "they see humor when no one else does. . .they see the depth of things." Thus, Kitty goes beyond literal comprehension in her discussion of literature. Classroom observations support this. In fact, Kitty is one of those individuals for whom questions and follow up questions come naturally. She intuitively provides scaffolding for some students to build a concept while withholding support from those she expects to construct meaning independently.

Finally, she demonstrates an understanding that comprehension is mediated by word recognition. "Their comprehension when read to was fairly good. . .but they don't know the words." By reading to students and then discussing the content, she knows that she can measure comprehension only. This information would then be considered when a student did poorly on a comprehension worksheet. She would then inspect the readability of the sheet for further insight into whether the student needed reteaching in comprehension or word recognition.

Oral language. Her responses in the area of oral language are representative of other subjects. She states that oral responses reveal concept development and comprehension. She feels that choice of words and oral vocabulary often reveal whether a student is a fluent and eager reader. She is particularly observant of students with articulation difficulties as she feels this affects mastery of phonics skills. She explains about a student:

. . .She has very few phonics skills. . .she doesn't because she has a terrible speech problem. . .and she doesn't hear the sounds. . .

Kitty demonstrates an understanding that knowledge of word meanings in reading proceed from a child's receptive vocabulary. If
a word is not in students' receptive vocabularies, they may decode it but will not comprehend it unless they are able to construct enough meaning from the surrounding context. This concept is expressed in the following quotation:

\[\ldots\text{C. has the sounds but he can't blend them.} \ldots\text{he can say the sounds in isolation.} \ldots\text{but he cannot blend them and come up with a word.} \ldots\text{and I don't know if it's because C. doesn't have the vocabulary and it doesn't click in that when I put these things together.} \ldots\text{this is making a real word.} \ldots\text{but he'll come close every time.} \ldots\text{but he never.} \ldots\]

Finally, Kitty describes the oral language of above average readers. This theme has been expressed by other subjects:

\[\ldots\text{Oh.} \ldots\text{their choice of vocabulary.} \ldots\text{they see humor when no one else does.} \ldots\text{their oral expression.} \ldots\text{they see the depth of things.} \ldots\]

It is logical to infer here that "oral expression" means the elaboration of their oral responses and not word recognition fluency in context.

**Prior Educational data.** Kitty's comments in this area are representative of other respondents. She does glance at report cards, however, before she works with the students in the fall. Further, she admits that she is looking for the behavior and work habits sections of the report card:

\[\ldots\text{I kinda just glance through them.} \ldots\text{um.} \ldots\text{I look at the report card to see how they did last year.} \ldots\text{just kinda go through them.} \ldots\text{and get an idea.} \ldots\text{I'm looking} \ldots\]
at the right hand side. . .probably more than I'm looking
at the left hand side. . .work habits. . .behavior. . .

Additional cumulative folder information includes level of
parent's education and occupation, and whether the child attended
preschool. She also looks at the Metropolitan Readiness Test because
she states she has given this test and understands it.

The individual reading folder is also a storage place for prior
data. Kitty states, "all I do is I look at the front to find out how
quickly they went through the last book." This is especially
interesting in that she is not utilizing the basal. Nevertheless,
she is interested in their pacing through the basals the previous
year. She values any narrative comments, "If there's something
written by the reading specialist. . .if it's a new child in the
county. . .I'll look at that." However, she declares that she would
not look at word recognition inventories from first grade. "They're
not really helpful. . .they're old." She states that she does not
look at basal testing data until after she has worked with the child.

Finally, she states she would look at the previous report card
to check grades after she has completed the first quarter report card
and before she sends it home. It appears that her motivation is to
insure that she initiates parent-school communication if the second
grade report card is significantly different from the previous year.
She explains, . . ."just to make sure that I'm not sending home
something that they've never seen before". . .

Grades. Kitty does not grade initial attempts as stated before
and waits to collect formative data until she perceives the majority
of students are ready. She states:

. . .When you introduce it, you work on it. When I feel
that the majority of them are comfortable with the skill
and have mastery of it, then I give a test, a worksheet,
whatever, a workbook page...

She states that she is uncomfortable with assessments that do not have "enough questions." She explains, "The reading scores are not based upon enough questions and therefore, I do not put much weight on them."

Her response to this need (to have comprehension grades that reflect a broad sample of behavior) is to grade often. She rationalizes the need to take many grades:

...To tell you the truth...for second graders I write grades daily...I take grades on comprehension...I take a grade on oral reading...and I take a grade on reading skills...

Kitty has already stated that she values group interactive performance over paper-pencil performance. She returns to this theme and appears to wonder aloud how she can reliably include these data in order for her grades to reflect what she feels is important:

...also there is just their daily performance...and I don't really know how to write that down...

She then provides a specific example where assessment of learning is more valid when accomplished orally:

...They write a really good story...but they can tell me so much more about the story they have read than they can put on the paper...

Kitty appears to recognize that these interactive and observational data are an important component of this child's
acquisition of literacy. At present they are not systematically recorded and are not reflected in the child's grade. Kitty's comments reveal the frustration she feels over this incongruity between her beliefs and her practices.

Kitty states that she does not favor letter grades because she does not think they present an accurate picture. However, because she has to give grades and be accountable to parents, her response is to collect many grades. She justifies this:

...I need a big enough pool to come up with a grade...
I'm not entirely comfortable saying to the parents, this child is a B student...this child is a C student...
even when I have the grades to back up this judgment, I don't think that gives the parents a good picture...

Grades appear to reflect learning and effort in this teacher's classroom. Kitty admits that she finds it very difficult to give low grades to students who are trying but experiencing difficulty. She talks about these students:

...I get really, really frustrated when I have a child like that...we worked with them...we've done everything we could...and they were trying really hard...and still they were making really, really poor grades. I try hard not to put their grades on their papers. I communicate bad news to parents and show them the work...and I have been known to give those children C's on their report card...

When papers provided for this subject were examined for grades, below grade level students were given comments such as "two out of five" written at the top.
Written language data. Kitty has stated limitations of generated written responses as a measure of comprehension. However, she feels they are a rich source of information that reflect back upon reading achievement. For example, she states, "their writing gives information . . .their sequencing . . .their expression of ideas. . ." In another interview, she adds:

. . .A lot of times you can tell how much they're reading . . .the way they express themselves . . .the length of their sentences . . .the choice of their words . . .the descriptors that they use . . .the punctuation . . .and by this time of year, you can really tell that they have been reading so much because their writing really shows it . . .

The following writing sample was provided by this informant. She did not comment upon it in any way except to say that they had just finished the book. There were no corrections on the paper. The word "wonderful!" was written across the paper:

Willber took good care of Charlotte's egg sac. Then one spring day the eggs hathed. All the little spiriders stayed for a few days and then most of them left. Willber thought all of them left when he saw three little spiredes and the names of the little spiredes are: Toy, Aranea and Nellie. Willber loved Charlotte's kids. Willber never forgot Charlotte. Charlotte is permetle stuk in Willber's heart. This is a story of friendship, sadness and caring. Right before Charlotte dieaed she said, "What is life anyway we are born, we love a little while, and we die." I think that was a very truthful
thing to say. Charlotte ment alot to Willber. (note: all punctuation and spelling errors are intact above.)

Unique factors affecting assessment context. There is only one unique factor affecting this subject. However, it is not discussed here to avoid including identifiable data.

Influences upon development of assessment expertise. Kitty readily admits that she did not begin teaching with many assessment strategies. As stated before, her first year she could identify only those at either end of the learning spectrum in her classroom. She laughs and says, "Oh, my gosh! I was surprised at the end of the year when they could read. I was really shocked!"

As many other subjects, her growth occurred as a result of trial and error. When asked how she developed her skills, she replied:

. . just by trying all kinds of different things. .
.finding out what worked with a child. . .what doesn't work. . .making mistakes. . .saying, "oh no," and doing it again. . .

She appears to combine the stance of those teachers who look for patterns of behavior to assist them in diagnosing learning needs, and those teachers who say each class is unique. She summarizes her position:

. . .You do think you've seen this before. . .it's a pattern you do. . .but you do that with everything. .
.most kinds of patterns I think are just part of us. .
.we're always saying, oh well, if so and so does this, they're going to have problems. . .but I really believe
that if we didn't do that... see patterns... every
class would be new... and every class is different...
and you start off doing some things... but I think that
through the years you develop so many different ways of
approaching reading, phonics, writing, whatever it is...
and sometimes it works and sometimes it doesn't... but
you just keep trying...

Kitty identifies a principal and a reading specialist as
contributing to her development as an assessor. When asked what she
learned from the principal, she states:

... the thing that sticks out in my mind the most is
that she knew the whole child... and everything about
that child... and I think she made me more aware of just
really looking at the child, thinking about the child and
everything that he does... trying to put it all
together... I was just always fascinated that I could
stop her in the hall and say, I'm having a problem with
thus and so, and she could just... tell me more about
the child...

When asked what strategies she uses to organize and value all
of these data, she laughed and said, "Cross your fingers and pray!"
This humorous response may be an indication that while Kitty knows
that multiple sources of data are better than a single measure, she
does not have a definite plan for collecting and valuing data
deliberately for decision making. In addition, she does not appear to
decide apriori what assessment data to collect in order to measure a
specific learning target. Specification of learning targets,
deciding upon the purpose of assessment, and selection of appropriate assessment data are reflected in authentic assessment literature as current best practice (Stiggins, 1994).

When asked what assessment strategies she had learned from the reading specialist, she states:

...There are so many little things that I've picked up from watching her with children...just observing...

...just one more piece of information...looking at their writing...using the cloze procedure...listening to them read...

Concerns and recommendations. When asked what her concerns were in light of the transition of reading and assessment, she states:

...that some child doesn't slip through the cracks...

...that's the thing that really, really frightens me...

...the quiet child...the child that you really can't get a good handle on...the one who might sit there with a book...who might sit there doing something...and have needs that you don't know about...

Her response to this concern is to make time to spend with each child and to continue to "go back" to be certain she has an accurate picture of each child's performance. She continues:

...You just have to...some of those other things...

...I think just make a real conscious effort to note that this child is going to be quiet...this child is not going to cause any problems...he's not going to stand out...and just to make time that you keep checking...checking back...
Kitty's further concern in the area of authentic assessment is the time required to collect data:

... even putting together the portfolios. . .I think are wonderful. . .and I'm really excited about that. . .but we've got to have time to spend with the children every day. . .or I do not feel comfortable about it. . .

Her first response when asked about new teachers and their training needs was, "I feel sorry for them!" On a more serious level, she declares that it takes experience in order to develop a schema of what a second grader can do and should do:

. . .It takes experience. I really think that it takes several years of working with children and knowing exactly what a second grader can do. . .or knowing what a child on that grade level can do and should be able to do. I don't think you can pick it up in student teaching. You can have an idea. I don't think you can read about it and really know until you have worked with them. . .

She makes several recommendations to improve teacher inservice training. She explains:

. . .Observing. . .I thinking observing. . .I think they really need to spend time with the older teachers. . .more experienced teachers. . .mentoring. . .working together. . .having that teacher come into their classroom. . .having the experienced teacher go into the
new teacher's classroom. . .and just listen to the
children. . .talk with the teacher about what she sees. .
.and what they're doing. . .

This is a recurrent theme: the mentor needs to be in the new
teacher's classroom in order that the novice can attempt to interpret
data and the mentor can elaborate on this understanding. This
follows the constructivist model of learning. The mentor assists the
new teacher in building a schema for receiving and valuing bits of
data that float around classrooms.

Written documents provided. This subject provided two sets of
papers. The set of papers for the below grade level group was a
cloze activity from the board with reading vocabulary. Errors were
circled and the number right out of five was written at the top.
Grades included three perfect papers, a three out of five, and a two
out of five. In addition, reading response logs for Charlotte's Web
were provided. No corrections were made on these papers. One paper
had "wonderful!" written across it. The think aloud was a
metacognitive activity that has been previously discussed.

The observation was an instructional activity with Charlotte's Web. The teacher read aloud while others followed along. Three
children in the back of the room did an alternative activity. When
children were allowed to select partners, Kitty chose three children
to read with her. Follow up questioning revealed that she felt their
oral reading needs were greater than the rest of the group, and she
wanted them with her to avoid frustration. Therefore, although this
was a whole group activity, modifications were made for six students.
**Cross-Case Analyses**

**Introduction**

Tables 9 to 16 provide information about the data categories that emerge as most prominent across all profiles and the percentage of reliance upon those data sources by the subjects. This section is organized around those categories of data. Percentage of reliance is reported along with descriptive statistics for each category. Explanations for reliance upon a category are provided by text chunks from interview data and discussion. Extraneous influences upon the data are presented.

Following this, data are then grouped according to emerging themes. Thus, data are presented as aggregates for decision making. A pie graph is used to display the aggregated information in Graph 1.

In addition, themes that have emerged during data analysis and interpretation are discussed. These results are compared with preceding theory reviewed in Chapter 2.

A data profile is constructed from common elements of the seven teacher profiles presented in this study. Common uses of data are discussed.

Finally, the resulting profile which emerged from this study is compared with a constructivist model of literacy and authentic assessment practices. Discrepancies between the constructivist model of literacy and teachers' practices are noted.

**Word recognition.** Interview data reveals that all informants reported word recognition as their primary source of data for decision making in reading with the exception of Betty for whom it was the third most important data source. A more fine-grained analysis reveals that the greater frequency of these responses refer to oral reading.
<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitty</td>
<td>21%</td>
</tr>
<tr>
<td>Stacy</td>
<td>18%</td>
</tr>
<tr>
<td>Roberta</td>
<td>15%</td>
</tr>
<tr>
<td>Amy</td>
<td>12%</td>
</tr>
<tr>
<td>Scotty</td>
<td>11%</td>
</tr>
<tr>
<td>Elaine</td>
<td>11%</td>
</tr>
<tr>
<td>Betty</td>
<td>8%</td>
</tr>
<tr>
<td>Mean</td>
<td>14%</td>
</tr>
<tr>
<td>S.D.</td>
<td>5%</td>
</tr>
</tbody>
</table>
Roberta reported 33 oral reading responses and eight phonics related responses while Kitty reported 17 for oral reading, nine phonics, and nine related to analysis of miscues. Clearly, word identification data heavily influence decision making by these teachers. Kitty is representative when asked how she sizes up a group in the fall, she responded, "just by spending time with them, talking to them, and listening to them read."

It should also be noted that respondents were often unclear about what they meant by "know vocabulary." At times they were referring to decoding or word identification and at other times they referred to knowledge of word meanings, or the ability to use a word correctly in a sentence orally. As this ambiguity emerged, teachers were asked to clarify the context of their statements.

It should be noted that the prominence of word recognition data may be a function of the level of concern these teachers hold for three or four students in their class who display severe needs in this area. When teachers' responses were probed, they all indicated that they do not collect the same type of word recognition data on all students. Kitty says, "I don't do flash cards with the other group at all."

However, when asked to describe their reading program and select students who are learning or not learning and describe how they know this information, teachers were much more likely to talk about students who concerned them. Often these were students whom the teachers had not yet figured out diagnostically. Therefore, their interview responses resemble an attempt to make sense of all the data they have on a particular student and reach a conclusion. It is possible, therefore, that these results are skewed by teachers' tendencies to return to those students who present severe decoding needs. Semi-structured interview questions allowed teachers great
laterality in selecting examples to illustrate their answers.

The prominence of word recognition data is consistent, however, with the findings of previous data studies. Barry (1992) surveyed 206 teachers and found they favored oral reading data. Pryor (1992) found that the three first grade teachers she studied emphasized letters and sounds as their primary data source. In addition, Antonacci (as reported by Pryor, 1992) found that kindergarten and first grade teachers assessed below grade level students more often. These assessments were focused around themes of decoding also.

The seven teachers in this study appeared to lack an organized method for recording, valuing, and using word recognition data in their assessments. Kitty and Amy report grading oral reading but fail to articulate criteria to students. Stacy records words missed and puts the list in the student's reading folder. She then retests for mastery. She and Kitty report recording errors phonetically. Stacy states she does this for her low group and then specifically states they have difficulty with wh words. Kitty records these errors for one student to share with the learning disabilities specialist. Betty mentions recording how a child reads to provide documentation at a parent conference. Field notes reveal, however, that this documentation was a placement IRI performed by the reading specialist. Betty does not conduct IRI's or running records herself.

Reliance upon memory for documentation of word recognition data appears to be the strategy used most often. This is consistent with the findings of Stiggins and Bridgeford (1985), Stiggins and Conklin (1992), and Pryor (1992). The latter reports the studies of Church (1990) and Nelson (1990) that concur. Betty is articulate and confident in describing how she chunks, stores, and retrieves this information. She relates that she keeps information about each student in her head like "little drawers." Each student has a drawer and she states the students are so different that she is able to
remember "basically what each one can read."

Word recognition data appear to be part of the interactive data base that teachers use to form their attributions about who is learning what has been taught. In this respect, it is a triangulation point for paper-pencil data such as basal tests and work samples. Should silent comprehension scores be discrepant from their attributions for a student's word identification accuracy, these teachers would collect more data and observe in an effort to resolve this inconsistency.

Although other data studies have documented primary teachers' orientation toward word recognition data (Stiggins and Bridgeford, 1985; Pryor, 1992; Stiggins and Conklin, 1992) constructivist models of reading (Goodman, 1989; Harp, 1993) only acknowledge word recognition as a window for looking inside the reading process to gain an understanding of how a student is interacting with the text. Therefore, the collection of these data appears to have little importance by itself. It is only with the interpretation of the experienced practitioner and observer who understands the context of the measurement as well as the knowledge the child brings to this text, that the measurement takes on importance in describing a child's literacy development. Stacy and Kitty appear as outliers in offering more error analysis than other subjects. It should be noted, however, that Kitty alone attempts to understand students' strategies to construct meaning while reading and in understanding the importance of word recognition in the larger context of reading as a constructive process.

Comparison of data. Next to word recognition data, comparison of data was the category with the highest percentage of responses. In fact, this category was not in the original design. It was initially considered to be a strategy for decision making rather than a data source. However, in coding responses, it became clear that
data that had been triangulated were more highly valued; therefore, frequencies of this category were added to the data analysis.
Table 10

Percentages of Reliance Upon Comparison of Data

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betty</td>
<td>12%</td>
</tr>
<tr>
<td>Amy</td>
<td>9%</td>
</tr>
<tr>
<td>Scotty</td>
<td>9%</td>
</tr>
<tr>
<td>Elaine</td>
<td>8%</td>
</tr>
<tr>
<td>Stacy</td>
<td>8%</td>
</tr>
<tr>
<td>Kitty</td>
<td>7%</td>
</tr>
<tr>
<td>Roberta</td>
<td>7%</td>
</tr>
<tr>
<td>Mean</td>
<td>9%</td>
</tr>
<tr>
<td>S.D.</td>
<td>2%</td>
</tr>
</tbody>
</table>
The truncated distribution of these seven teachers on this data source of triangulated data appears to suggest it is a common and stable data source used for decision making in reading. The purpose for comparing data appears to be in order to gain confidence in decisions relating to grading. As with word recognition data, there does not appear to be any prior plan to compare data sources. This strategy is simply employed as discrepancies are noted in an individual's profile.

These primary teachers clearly stated that they "need a lot of data," "one more piece of information," and that they were reluctant to rely heavily upon one data source. Roberta stated, "...now test scores...grades...are very important...but they're not the whole story of a child..." Kitty recalled an important lesson learned from a previous principal. "I think she made me more aware of just really looking at the child, thinking about the child and everything that he does...trying to put it all together..." Betty and Scotty spoke of caution in placing confidence in written work of students who are distractible or immature. By comparing these students' written products to their oral responses, their oral language, their estimate of ability and oral reading fluency, they were able to deduce that some written products were not valid representations of what a student was capable of doing. Kitty issues a caveat regarding relying solely upon literature responses for evaluating comprehension. "They can write a story...a really nice story...but they can tell me so much more of what they learned from what they had read."

Comparison of data also included the strategy of comparing to a standard. This could be comparison to a baseline set by that child's previous performance. All subjects indicated they would consult previous report cards either before they met the class, after working
with them for a while, when they noticed difficulties, or just before sending home the first report card. They justify their need for this information because they want to be certain that if their assessment of the child differs from the previous year, that they initiate communication with parents prior to the report card. Comparison to a child’s previous performance, therefore, appears to be concerned with growth and accountability.

Secondly, teachers mentioned comparing a child’s performance to a standard of what second graders should be able to do. Kitty declares that one develops this standard through experience only. Amy states that substituting in a variety of grades made her aware of these standards. This comparison appears concerned with comparing and sorting as well as with accountability. Their intuitive understanding of what constitutes grade level expectations appears to be normative.

There are several possible explanations for teachers’ concerns with some type of normative standard for grade level work. First of all, this may be due in part to a basal driven curriculum. Secondly, it is important to understand the lack of consensus within the State of Virginia in the area of grade level expectations. The state department of education and local school divisions have published grade level learning objectives. However, these learning objectives do not indicate performance standards. Thus, it appears that grade level performance standards are not publicly articulated and are conceived for the most part in the minds of teachers. The benchmark used most often in the absence of performance standards, therefore, is the level of proficiency required by either the Virginia Standards of Learning Test or the basal test for a grade level.

A third standard for comparison is the group’s performance. Dahlof and Lundgren as reported in Stiggins and Conklin (1992) report the use of a "steering group." Thus, teachers identify a subset
within the classroom or within an instructional group. They then monitor the cues of these students to determine their mastery. Teachers then pace the group based upon the needs of this chosen subset. This comparison appears concerned with instructional monitoring and revision.

The teachers in this study also spoke of a group achievement standard and spoke of children who were too advanced or slow for a given group and required regrouping. Kitty talks about the fact that only one student in her low group appeared to profit from phonics instruction. Regarding comprehension, they speak of children who "get more from the story." In writing they refer to those who have more elaborated language and word choices as exceeding the class or group standard.

Observational data. Observation as a data source varied widely among the seven teachers from two percent (Elaine) to 13% (Betty).
Table 11

Percentages of Reliance Upon Observational Data

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betty</td>
<td>13%</td>
</tr>
<tr>
<td>Roberta</td>
<td>10%</td>
</tr>
<tr>
<td>Stacy</td>
<td>9%</td>
</tr>
<tr>
<td>Kitty</td>
<td>6%</td>
</tr>
<tr>
<td>Scotty</td>
<td>5%</td>
</tr>
<tr>
<td>Amy</td>
<td>4%</td>
</tr>
<tr>
<td>Elaine</td>
<td>2%</td>
</tr>
<tr>
<td>Mean</td>
<td>7%</td>
</tr>
<tr>
<td>S.D.</td>
<td>4%</td>
</tr>
</tbody>
</table>
Betty's responses were replete with references to children's body language and her attributions for their meaning:

...the boy in the red shirt. . .if he's sure of the answer, he'll keep saying it over and over. . .if he's not, he'll let the others have their say. . .I can tell by how you sit, how you write your name on your paper, if you're angry with me. . .either you're writing is so small or it's so super, super sloppy that I can't read it. . .all of those clues tell me what kind of day you're having. . .why you're not able to read. . .

Although it was not possible to judge the accuracy of these comments, it was clear that she was indeed a careful observer.

Reliance upon observation is well documented in the literature from the work of Dorr-Bremme and Herman (1986), Stiggins and Bridgeford (1985), and Pryor (1992). It is possible that the creation of categories of data within this study for oral responses, word recognition, oral language, behavior and work habits reduced the apparent reliance upon observational data. Therefore, these categories of data are aggregated in Table 12 to produce the category of interactive data. The rule used to aggregate data is that none of these sources of data require a paper-pencil response from the student. Therefore, observation is the only strategy available to capture this student information as the performance leaves no product or evidence.

Interactive data. All teachers in this study rely upon this source of data for over half of their data base to inform decisions in reading. This study, therefore, adds to the case study literature in this area. It is of interest to note that Kitty displays the
strongest orientation toward whole language instructional methodology. She also appears to rely most heavily upon interactive data.
Table 12

**Percentage of Reliance Upon Interactive Data**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitty</td>
<td>75%</td>
</tr>
<tr>
<td>Stacy</td>
<td>72%</td>
</tr>
<tr>
<td>Amy</td>
<td>69%</td>
</tr>
<tr>
<td>Elaine</td>
<td>64%</td>
</tr>
<tr>
<td>Scotty</td>
<td>59%</td>
</tr>
<tr>
<td>Roberta</td>
<td>59%</td>
</tr>
<tr>
<td>Betty</td>
<td>51%</td>
</tr>
<tr>
<td>Mean</td>
<td>64%</td>
</tr>
<tr>
<td>S.D.</td>
<td>8%</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
While interactive data is the largest aggregate of data reported in this study, it is largely undocumented. In addition, it is not used for grading. As stated, its most significant use for these seven teachers appears to be as a base for their attributions and as a triangulation point for written data. Although this data source is much richer for information about the student's interaction with text, it is generally not reported to parents and is not available for later reflection. Finally, there is a great potential for error as this data is highly valued but not systematically recorded.

**Basal test data.** Reliance upon data from basal tests also varied from three percent to ten percent.
Table 13

Percentages of Reliance on Basal Test Data

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roberta</td>
<td>10%</td>
</tr>
<tr>
<td>Elaine</td>
<td>8%</td>
</tr>
<tr>
<td>Betty</td>
<td>7%</td>
</tr>
<tr>
<td>Scotty</td>
<td>6%</td>
</tr>
<tr>
<td>Stacy</td>
<td>4%</td>
</tr>
<tr>
<td>Amy</td>
<td>3%</td>
</tr>
<tr>
<td>Kitty</td>
<td>3%</td>
</tr>
<tr>
<td>Mean</td>
<td>6%</td>
</tr>
<tr>
<td>S.D.</td>
<td>3%</td>
</tr>
</tbody>
</table>
Responses in this category tended to parallel the teacher's progress in transitioning to a multi-text whole language program. In fact, it should be noted that Amy, who relied the least on this category of data, also reported four percent of her data from performance assessments or checklists. Roberta stated that she "grades everything." In addition, she chooses to design and grade activities that she knows are tested on the basal test. Therefore, it is not surprising that her reliance upon this source of data is the highest among subjects. It should be noted that although the percentage of reliance is not as high for this data source as for interactive data, it may have more influence over decisions in the area of accountability because it is documented, unlike information obtained from interactive data sources. This will be discussed later under types of decisions.

It appears that data collection may be a function of instructional methodology and external mandates in this area. Text segments document that because teachers are held accountable for the results of basal tests, they will allot instructional time for tested skills and they will collect formative data to insure mastery.

Behavior and work habits. Behavior and work habits as a data source varied from four percent to eleven percent.
Table 14

Percentages of Reliance Upon Behavior and Work Habits Data

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scotty</td>
<td>11%</td>
</tr>
<tr>
<td>Betty</td>
<td>8%</td>
</tr>
<tr>
<td>Elaine</td>
<td>8%</td>
</tr>
<tr>
<td>Amy</td>
<td>7%</td>
</tr>
<tr>
<td>Roberta</td>
<td>6%</td>
</tr>
<tr>
<td>Kitty</td>
<td>5%</td>
</tr>
<tr>
<td>Stacy</td>
<td>4%</td>
</tr>
<tr>
<td>Mean</td>
<td>7%</td>
</tr>
<tr>
<td>S.D.</td>
<td>2%</td>
</tr>
</tbody>
</table>
These results are interesting in that Stacy reported having a mainstreamed student with an identified disabling condition. She then clearly discriminated between this student's behavioral needs and his reading needs, stating that:

... behaviorally... emotionally he's different from everyone else but as far as academically... when you see this child... as long as he's actively engaged in an activity, he's usually pretty controllable...

It should not be inferred that other subjects devoted a disproportionate amount of time on management issues. Many comments deal with learning cues from student behavior rather than disruptive behavior that must be managed. Betty states, "I can tell by the way you write your name on your paper, what kind of day you're having."

Oral responses. Oral response data appear to be consistent across all profiles.
Table 15

Percentages of Reliance upon Oral Response Data

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roberta</td>
<td>9%</td>
</tr>
<tr>
<td>Amy</td>
<td>7%</td>
</tr>
<tr>
<td>Elaine</td>
<td>7%</td>
</tr>
<tr>
<td>Scotty</td>
<td>5%</td>
</tr>
<tr>
<td>Stacy</td>
<td>5%</td>
</tr>
<tr>
<td>Betty</td>
<td>5%</td>
</tr>
<tr>
<td>Kitty</td>
<td>4%</td>
</tr>
<tr>
<td>Mean</td>
<td>6%</td>
</tr>
<tr>
<td>S.D.</td>
<td>2%</td>
</tr>
</tbody>
</table>
Teachers generally concur that in developing assessment strategies during their career, they have moved from reliance upon paper and pencil products to reliance upon interactive data gained during teaching. Kitty stated, "I can just tell by his responses. He doesn't get the main idea of what we are discussing. He'll give responses like 'the pig was little.' He doesn't grasp the meaning." Teachers also uniformly reported relying upon memory to document this data source. Betty summarized with her statement:

"...dealing with these four students and I just keep it in my head...there's little drawers in my head...you know...like the Mickey Mouse Club...this is M.'s drawer...and this is A.'s and this is S.'s, and D.'s...and you know basically who can read what..."

It also appears that these teachers derive cues from the quality of students' oral responses regarding how much scaffolding to provide to assist students in building meaning as they read. A classroom observation in Stacy's room during a lesson on word meanings revealed that she responded with verbal praise to a correct response, while she repeated and elaborated upon answers that were approximate.

When confronted with an incorrect response, however, she accepted the part of the concept that was offered and continue to question other students in order to assist students in piecing together the meaning of the word family. Zemelman, Daniels and Hyde (1993), in writing on best practices in reading, feel this scaffolding can be deliberately built into classrooms. They compare the hypothesis testing strategies of acquiring language to those of acquiring literacy. Children must be allowed to make errors as they construct meaning.
Oral language. Oral Language was also added as a data category in the process of coding text chunks, when it became obvious that informants were drawing a distinction between oral answers in class and the overall quality of a student's oral language. Teachers reported gleaning much information from oral language including an estimate of general ability, prior knowledge and background of experiences, and knowledge of concepts.
Table 16

Percentages of Reliance upon Oral Language Data

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amy</td>
<td>9%</td>
</tr>
<tr>
<td>Scotty</td>
<td>8%</td>
</tr>
<tr>
<td>Roberta</td>
<td>8%</td>
</tr>
<tr>
<td>Kitty</td>
<td>5%</td>
</tr>
<tr>
<td>Elaine</td>
<td>4%</td>
</tr>
<tr>
<td>Stacy</td>
<td>3%</td>
</tr>
<tr>
<td>Betty</td>
<td>3%</td>
</tr>
<tr>
<td>Mean</td>
<td>6%</td>
</tr>
<tr>
<td>S.D.</td>
<td>2%</td>
</tr>
</tbody>
</table>
Oral language was frequently reported as a source of data for comparison. Teachers reported comparing oral language to oral reading fluency, or the quality of oral language to the quality of written expression. If they noted a severe discrepancy in the two, their response was to gather more data. Their reported data gathering activities were to observe the child more closely, ask the reading specialist to test the child, seek information from the parent, talk to the previous teacher, or refer the child to the Child Study or Special Education Committees.

The findings of this study with regard to oral language appear consistent with the expectancy literature. Teachers predict a student's potential to achieve from a variety of cognitive and noncognitive cues (Calderhead, 1983; Salmon-Cox, 1981; Gaines and Davis, 1990). Oral language appears to be an important cue for these teachers. In addition, their expectancies appear global as evidenced by their interview responses such as "he's a bright little boy but something is not clicking." All of these teachers report revisions in their groupings and interpretations as the year progresses. This suggests that their continued data collection does produce changes in their expectancies.

**Comprehension.** Percentages of reliance for comprehension are not listed because of the low frequencies reported. In addition, some comprehension data may have been unrecorded and coded under basal test data, oral responses, or work samples, as respondents did not specify the data as comprehension.

When asked to describe their philosophy of reading, informants concurred on a definition that included reading for meaning, reading independently, and reading for enjoyment. The parameters of these definitions suggest an emphasis upon comprehension in data gathering. However, as stated before, when teachers were asked to describe
students' growth as readers, they chose students with decoding needs and focused their comments on efforts to improve their word identification strategies in isolation and in context.

It is also possible that both first and second grade instructional programs focus upon learning to read, which is replaced by reading to learn in the higher grades. A replication of this study with elementary teachers in grade four of five would yield useful data in an attempt to understand the preoccupation of primary teachers with word recognition data.

An analysis of teachers' actual comprehension statements, videotapings of comprehension discussions, and classroom observations, however, reveal that these teachers do not focus their questioning primarily on factual recall as reported by Stiggins, Griswold, and Wikelund (1989). Indeed, if an area of comprehension predominates, it is that of concept building for word meanings such as hero. Students were encouraged to provide partial meanings. The concept was constructed from these bits of concepts, responses to scaffolding questions, and the elaboration provided by the teacher. Secondly, main idea of the story was most frequently mentioned in interview data.

Accountability. One common theme among informants to emerge from the data in this study was their need for data that would enable them to describe the achievement and growth of their students accurately to parents and administrators. They expressed concerns about articulating clear performance targets. Betty stated emphatically, "just tell me where you want me to be, where you want me to go...but let me get there my way...but I need direction."

They also shared the common belief that performance assessments or portfolios must be characterized by some level of standardization to preserve the integrity of the assessment, and to insure uniformity in administration, scoring and interpretation. They described the
basal program as providing structure both in methodology and assessment. While eager to let go of the basal content, they expressed concerns about abandoning the assessment component. The role of phonological awareness and measurement of sight vocabulary as well as reading fluency were recurring themes.

Stacy was unique, however, in her statement about decision making with performance assessment data. She stated her awareness that performance data often lack the decision making parameters that teachers have been accustomed to with criterion referenced instruments. "...oh yeah, you could test from here to June 18, and not know what to do with it...so yeah, you have to know what it's going to show you." She alone appeared to understand that performance data will require teachers to collect data deliberately to answer their questions, interpret and value the data, make decisions about instructional revision, and interpret both the data and decisions to parents and other stakeholders.

**Response to changes in assessment.** All subjects had some experience in performance assessment. The new district math curriculum includes options for either paper-pencil or performance assessment. Those who had used the latter spoke at length about the time involved in one-on-one testing and how much instructional time they perceive this type of assessment takes. They made no comments about the quality of the information obtained from the two types of testing.

When asked about changes in reading assessment, Betty was adamant that teachers did not have the time to develop these instruments. She extends this discussion to suggest that if teachers develop different performance assessments on their own, they may not be equally valued by teachers in succeeding years.

**External mandates and assessment change.** The findings of this study echo those of Pryor (1992): that teachers are provided with
inservice and training in a new belief system about how literacy is acquired, but they are constrained by the external mandates from the previous literacy paradigm. These seven teachers concur that they would incorporate more whole language activities if they did not have to administer the basal test. This is even true in schools that adhere very loosely to the decision-making parameters of the criterion scores. Teachers feel if they must share these test results with others, they must prepare students to do their very best.

The purposes and uses of assessments. It is clear from the interview responses that these teachers view assessment and grading synonymously. Stiggins (1994), however, distinguishes between assessment as the collection of information about students and grading as the process of abstracting a great deal of information into a single symbol for ease of communication. The belief systems that underlie the grading strategies of these subjects appear to differ greatly. Roberta appears to be at one end of the continuum where she "grades everything." Amy, however, is very reluctant to assess children at all for fear that they will not be successful as she states, "I wouldn't dare until they're ready." There are commonalities, however, among subjects. Six of the seven concur that initial attempts at a skill should not be graded. They also articulate that they do not record grades for cooperative papers and those completed with assistance.

They do display an understanding of options for using grades for motivation, however. Betty grades handwriting when she perceives it is necessary to motivate students to be neater. Teachers vary on whether they record oral reading grades; however, they concur in not revealing to students what the criteria are for oral reading grades. This appears to be motivated by a concern for students' self-esteem. If pupils do not know they are being graded, they will not be
concerned about the results. This is not consistent, however, with recommended best practices (Stiggins, 1994).

Subjects admit factoring noncognitive behaviors into their grading strategies. They factor effort, motivation, behavior, and work completion into their grades. When talking about weak readers who try very hard, Kitty states, "I have been known to give them a C."

There appear to be no indications throughout all interviews that teachers differentiate between the assessment functions of determining mastery and instructional revision, with those of accountability and program evaluation. Teachers certainly feel accountable as evidenced by the comments, "I've got to have data to show them," or "I needed [this basal test] to prove they were as good as I said they were."

There are several discrepancies between the reported practices of these teachers and recommended best practices from authentic assessment literature. Stiggins and Conklin (1992) and Harp (1993) advocate clear criteria before collecting data that is used for accountability purposes. Further, students should be made aware of these criteria. Factoring such noncognitive variables as ability, motivation, and work habits into grading schemes creates "measurement noise" according to Stiggins (1994). If one wishes grades to convey a consistent and accurate statement about learning, grades must reflect only measurement of learning outcomes. However, Stiggins (1994) argues that there is value in reporting this information, but not in grading it.

The experts agree that students should not be graded when the teacher is collecting diagnostic information prior to instruction, when a skill is first introduced, or when students are practicing skills and learning from their mistakes (Stiggins, 1994). However,
the experts do recommend that assessment information be collected throughout the learning process and that students be given feedback about their performance. This distinction is impossible to make if assessment and grading are viewed as one and the same and if teachers do not differentiate among the different purposes for assessment. Roberta, who "grades everything," and Amy, who waits to grade until students have mastered the skill, exemplify the extreme positions of failing to distinguish among assessment purposes and the uses of grades. By failing to make these distinctions, assessment information and resulting grades contain "noise" that makes interpreting grades or making decisions based upon them difficult or unreliable. It is clear from this study that training in classroom assessment purposes, methodology options, and grading and reporting methods is warranted.

Sizing-up strategies. The results in this area are not surprising, but this study does contribute a fairly clear picture of what data sources teachers consider in making fall grouping decisions. Previous studies suggest that once these decisions are made, they remain fairly stable throughout the year (Doherty and Conolly, 1985). Therefore, it is important to understand what informs these judgments.

Sizing-up decisions appear to be based on inspecting a writing sample, listening to oral reading, talking to the student, asking them questions about a story, observing them in the classroom, looking at their pacing through texts the previous year, and attending to the behavior and work habits section of the report card from first grade. Oral reading, however, is mentioned most often in interview data as a sizing-up activity.

Typical Data Profile

Introduction. This section presents a typical data profile
drawn from the common elements of the seven teachers' profiles. It contributes one piece of the assessment puzzle to the literature. That is, it presents a composite portrait of the data sources and how they influence the reading decisions of these second grade teachers. It is important to remember that they are experienced teachers who are in a process of change from eighteen years of a basal driven reading program to a whole language program with authentic assessment practices. Therefore, the generalizations that may be drawn from this theory extend to experienced primary practitioners. Survey data (Barry, 1992; Coulter, 1992) and studies of teacher change (Borko, in press) suggest, however, that these teachers are very similar to those who are not early adopters of innovation. Barry reported the majority of teachers to be supplementing basals with whole language strategies. It is also likely that their belief systems and assessment strategies reflect both inductive and deductive approaches during the period of transition.

Thus, gaining a thorough understanding of these teachers may inform inservice providers and professors of graduate courses in this area. These experienced teachers are in a position to have the greatest effect upon preservice teachers as they move into the school environment for practicum observations and student teaching. If the knowledge base and belief system of preservice teachers differs qualitatively from the cooperating teachers with whom they train, these young teachers may be confused by the discrepancy. Many teacher training programs may be approaching reading and literacy development as a constructive process. However, novice teachers may perform their internship with teachers who are transitioning from a basal approach to a whole language approach. Therefore, novices may observe both constructive and discrete skill influences upon instruction as well as assessment.

Classroom assessment context. This classroom is characterized
by supplemented basal instruction. Although the basal is used, it may be used nonsequentially, thematically, or for whole class instruction. Workbook pages and copied sheets are available, but are rarely used. When they are given to students, it is usually for the purpose of assessment of a skill that has been taught and reviewed. One or more groups may be working on a novel and not using the basal at all during those weeks that the novel is instructed. Follow up work is likely to be journal entries, retellings, summaries, response to literature, or creative applications of the story. Children may read orally, listen to the teacher read, read silently, or read with a partner. Children are instructed in small reading groups or in large group settings.

Test data. Tests that accompany the basal are still given although there may be little connection between instruction and testing. That is, test results are not routinely used as monitoring tools to revise instructional placement, pacing, or the content of lessons. The tests may be given as summative assessments at the conclusion of units, at the conclusion of books, or at the ends of the semesters. Results are filed in the reading folders.

Sizing-up strategies. When students enter the class in the fall, the teacher gathers data on oral reading, writing, language, comprehension, and behavior. She may consult the cumulative folder for pertinent home background information and previous report card data. She may consult last year's teacher. She will double check the first report card in November with this previous report card for discrepancies. Should she discover a marked change in behavior or achievement from the previous year, she will initiate a conference with the parents to share this information and offer explanations before the report card goes home.

Word recognition data. Once groups are formed, word recognition data will continue to be collected regularly on those
students perceived to be functioning below grade level. Oral reading in context and in isolation will be measured as well as retention of previous vocabulary. Vocabulary lists will be provided to parents for reinforcement. Readers on grade level and above may have purposeful oral reading weekly, but the intent is not to collect diagnostic information.

**Comprehension data and oral language data.** Comprehension is monitored through oral discussion of oral or silent reading. Silent reading is considered to be more challenging because the teacher cannot monitor the process of meaning-getting and provide support where needed. Attention is paid to the correctness of the oral response as well as the quality of language the student displays. Criteria here include the length of utterance, grammatical agreement, choice of vocabulary words, and variety of syntax.

Based upon the perceived approximation of the oral response in terms of concept development, recall, thinking skills, and oral expression, the teacher will make a decision regarding the provision of scaffolding. That is, the teacher will provide assistance designed to move the student closer to a full understanding of the text read. This assistance could be additional questions, information provided, requesting information from other students, restating the student's response, etc. Wait time appears to vary directly with perceived difficulty with the concept presented. Those students perceived to be most independent are provided with greater wait time and less scaffolding. This appears to echo the results of expectancy research (Rosenthal and Jacobson, 1968). However, results here suggest that these teachers may be assessing interactively and providing instructional revision as the lesson continues. Students who appear to acquire a given concept with little difficulty receive less assistance in retrieving or formulating an oral response.
Provision of scaffolding appears to be a deliberate and/or intuitive response to a partial concept given by a student. This response by the teacher appears to be motivated by concerns for the student's self-esteem and a desire to provide just enough assistance to lead the student closer to the concept. Knowing how much assistance is required in a given instance for a specific student requires judgment regarding knowledge of the student's background of experiences, critical thinking skills, comprehension, affect, etc. as well as a thorough understanding of the mental processes required by the question under consideration. This is not the same as differential treatment according to ability as suggested by Rosenthal and Jacobson (1968).

**Written language data.** Information about reading development is also obtained through inspection of writing samples. Invented spellings are analyzed for information about phonic elements that have or have not been acquired, or to look at the effects of an articulation disorder upon the students' discrimination and encoding of phonemes. Other information gleaned from written language includes an awareness of the elaboration of the language, word choices, variety of sentence openers, sentence linkage, and correctness of punctuation.

**Pacing of instruction and interactive data.** Monitoring of the pace of instruction is provided by choosing a steering group within the larger group or the whole class. These students are considered average for the variance of this group. The teacher monitors them closely for understanding, and gauges the amount of practice provided and the rate of introduction of new skills upon their perceived mastery. Checking for understanding is done orally. Students who display partially correct answers are provided with scaffolding by the teacher to enable them to acquire the concept desired. Scaffolding may take the form of repetition of the student's
response, probing questions, or elaboration by the teacher.

**Grading and testing.** Grades may be taken on oral reading but are usually taken on worksheets after the group has been checked for understanding. Performance testing is rare, with the exception of testing the low group on flash cards of new or old vocabulary words or listening to students read a passage orally. The unit test is given and graded. Modifications are commonplace, however. These may include testing a student in a distraction-free environment, instructing a student to read the test aloud to himself, breaking the test into increments, instructing a student to reread a selection before reading questions, instructing a student to read the questions before he reads the selection, etc. Modifications are noted on the front of the test booklet.

**Decision making strategies.** Decisions are made very carefully by these teachers. It appears that decisions reported by the subjects in this study fall into two main areas. The first is concerned with instructional monitoring and revision. These decisions are informed by interactive data. These data include oral reading, oral responses, oral language, behavior, work habits, and observations. Performance assessments (where they are used) are grouped with interactive data because teachers use these as formative assessments only, typically do not systematically record information, and do not grade them. In a similar fashion, oral reading performance is seldom graded and is usually used for monitoring purposes only. Reliance on memory is used to store the information.

Attributions regarding a student's ability, achievement, motivation, behavior, and the influence of his/her home background are derived from this large pool of interactive data. Typical decisions made in this area include referral for testing or for a special program, or a change of instructional group or material. Other decisions include the pacing of instruction, methodological
decisions, and the amount of drill and practice provided.

Decisions regarding placement are typically shared with the reading specialist. This individual may also collect information such as informal reading inventory data. Interview responses suggest that these teachers regard IRI data as a rich source of information about a child's true performance level.

The second type of decision making is concerned with grading and issues of accountability. Written work samples, creative writing, and test data are collected to inform these decisions. However, the attributions formed from dynamic assessment activities appear to influence decisions in this area also. These teachers describe an elaborate system of error reduction strategies. That is, once they collect written products from students and grade them, they compare a student's performance to their attribution for his or her achievement. If a discrepancy exists, they collect more paper-pencil data, observe, interview, etc. They may discard pieces of written data that they subsequently judge to be unfair representations of a student's performance. Noncognitive factors may also affect grading decisions. These include effort and behavior. Finally, perceived ability may also mediate grading decisions.

**Aggregates of data used for decision making.** In order to gain some perspective on the uses of interactive vs. paper-pencil data for decision making, these categories of data have been aggregated and are displayed on Graph 1.
Figure 1
Aggregates of Data Sources

- Interactive Data
- Paper Pencil Data
- Prior Data
- Other

21.00%
11.00%
16.00%
52.00%
What is important to note here is that these teachers do rely upon interactive data for over half of their assessment information. This is a richer source of data than paper-pencil data, which account for only 16% of their information. Prior educational information accounts for an average of 11%. This aggregate includes standardized test data in the cumulative folders, notes in the reading folder regarding previous testing modifications, knowledge of a labeled status such as retainee, L.D., or Chapter I, home background information, information from the previous teacher, consideration of previous report cards, etc.
Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

In this chapter, conclusions are drawn from the study which are presented in light of recommended current practices in authentic assessment and a constructivist model of literacy. Based upon discrepancies between results of the study and current literature, recommendations are made for practice. Drawing upon an analysis of interview data, suggestions are made for preservice and novice teacher programs. Finally, directions for further research are considered.

Nature of Classroom Assessment Context

The classroom contexts studied in these seven classrooms were characterized by a high degree of complexity. This is consistent with preceding theory (Jett-Simpson, 1990; Pryor, 1992; Stiggins and Conklin, 1992). Teachers were confronted with vast amounts of data about students simultaneously. These included cognitive, social, and behavioral data. Teachers developed routines and procedures in order to process these data. These included reliance upon memory for documentation of all interactive data and grouping or chunking information mentally by student. In addition, the context for teachers interacting with perceived at-risk students appeared to be influenced by additional variables. These included the provision of varying amounts of scaffolding support to enable students to move toward an outcome, and the collection of additional achievement data on these students.

Because of the complexity of classroom assessment, implications for practice require ongoing efforts to provide novice and experienced practitioners with reliable strategies to select and
document appropriate data. American education cannot afford to wait for novices to acquire the experiential base required to process the multitude of data in modern classrooms. However, novices can be trained as observers and taught to use checklists to capture important interactive behavior for later reflection. In addition, reflection will help them understand the process of constructing meaning by their students. This will take initial training, on-site support, and continuing dialogue with caring, experienced practitioners.

The literature is replete with both recommendations and examples of instruments for data collection. Campione and Brown (1985) have labeled this process-oriented method of assessment as dynamic assessment, while Yetta Goodman (1978) first labeled it as "kid watching." It incorporates the belief that the classroom context for teaching and for assessing is an ever-changing one, and that multiple samples from this pool of information are needed to enhance the reliability of the data as a whole. Harp (1993) proposes that teacher intuition based on observation is as valid a measure for decision making as test scores. He quickly adds that it must be based on "careful observation and knowledge of a child's learning (p. 40)." Indeed, one of his principles of assessment is that teacher observation belongs in the center of the process.

**Teacher Profiles and Discrepancies with Current Literature**

The typical teacher profile of valuing assessment data that emerged in this study supports the fact that experienced practitioners rely upon the rich data sources of interactive teaching. They intuitively form attributions based upon these data. The data, however, are largely undocumented. When addressing questions of grading or accountability, the teachers abandon these data and attend instead to written products such as work samples,
basal tests, and creative writing. They attempt to discount invalid pieces of written data by constantly comparing students' written performances to the attributions they have formed about their students.

There are many instances in which the profile of data usage that emerged from this study does not conform to the findings of cognitive researchers such as Resnick and Resnick (1992) or Vygotsky (1978). First of all, it appears that teachers in this study focus their assessment strategies on verifying an achievement level or placement upon a continuum. This appears to conform more closely to the behaviorist theories of discrete skill accumulation. Once one has accumulated enough skills, one advances on the continuum. Cognitive theorists, instead, have provided educators with a model of interaction of text, reader, and context. The reader then adds his or her strategies in an effort to construct meaning (Jett-Simpson, 1990). It does not appear that data obtained from interviews in this study indicate that teachers collect assessment information to discover how students are constructing meaning from text.

There are many possible explanations for this. First of all, this school district has only recently begun the project of writing a reading curriculum. In the past 18 years, all learning targets were specified by the basal and measured through basal tests. All reference to learning targets associated with a given age or grade were similarly referenced from the basals designated for that grade level. Finally, promotion standards were tied to mastery of specific basal tests.

Therefore, the focus of instruction was upon mastery of content segments rather than mastery of specified learning targets. This is very similar to the findings of Clark and Peterson (1986) and Clark and Yinger (1979): that teachers plan in content chunks rather than to enable students to master learning outcomes. This is not
surprising when one considers that external mandates in reading also focus on content chunks.

Another discrepancy between the teacher profile that emerged from this study and recommended practice concerns the quality of data used for accountability decisions. Harp (1993) recommends that observation and teacher judgment be at the center of the assessment process. Goodman (1989) makes the point that teachers are evaluating children whenever they are engaged in interaction, observation, or analysis. The teachers in this study, however, appear to select data for important decisions that are tangible, easily quantifiable, and defensible. One subject even stated that she thought she was libel if she "made judgments about a child in writing."

It appears that several prerequisites are necessary in order for teachers in transition such as these to have the knowledge base as well as the freedom to become "kid watchers." These include:

1. The knowledge base of authentic assessment such as that expressed by Stiggins (1994), Harp (1993), Jett-Simpson (1990), and others. This would include specifying the questions(s) to be answered about a child's learning, deciding which behaviors might answer the question(s) posed, selecting methods of observing, documenting, and valuing these behaviors, and determining how to report these data obtained in an understandable form.

2. The knowledge base of the constructivist theory of reading. This would include the ability to recognize strategies students use as they interact with text, as well as the effect of context features on students' efforts to construct meaning. Teachers should also recognize the importance of what the reader brings to the text in terms of interest, prior experiences, language, decoding, concepts, and purpose.

3. The skills to document interactive classroom data. These might include the use of running records, anecdotal records, use of
checklists, reading logs, documentation of retellings, etc. This would also include skills in developing and using rubrics to judge performances and products.

4. The opportunity to participate in the process of formulating a reading curriculum that specifies learning targets by grade levels.

5. The revision of the external mandates for assessment in reading. Teachers, administrators, and policy makers should clearly distinguish between the purpose of assessment for external accountability and for monitoring student performance in the classroom. Selection of assessment methods and whether or not tograde would then suit those assessment purposes.

6. The availability of on-site support during the transition process. This would allow experienced and respected professionals to confront their basic belief system in a safe and nurturing environment. This in turn would facilitate integration of acquired knowledge, beliefs, and skills into the classroom context.

7. The participation of administrators in all phases of the transition.

**Preservice Teacher Training and Mentorship Programs**

All subjects reported that the greatest influence on their development of assessment expertise was an experienced teacher. The recommendations for training in this area that are a result of this study are listed below:

1. Preservice teachers should spend the first week of school in a classroom before student teaching. They should observe and talk with teachers about their thinking as they size up their classes and make managerial as well as planning decisions. This will provide novice teachers with the strategies needed to organize a new class.
Inherent in these organizational schema are strategies needed to organize a classroom in order to facilitate collection of interactive data.

2. Preservice teachers should be taught a collaborative model in teacher education courses. They should expect to work with an experienced professional for several years of their beginning teaching experience, adding to their knowledge base.

3. Mentor teachers should go into the novice teacher's classroom. They should observe students and the interactive data in this context. Finally, they should listen to novices think aloud about their assessment data and what they believe they have learned about their students' learning. In this way mentors will be able to offer insights in the most meaningful context, the teacher's own classroom. One of the major findings of Borko's study on facilitating change (in press) is that new information must be presented in a context where novices may see the applicability and understand where and how to integrate new concepts and beliefs into their daily practice.

4. Mentor teachers must become more reflective about data they process automatically. They must be able to describe clearly how they specify learning targets from the curriculum, as well as which data and what criteria inform their decision making. Mentor teachers must acquire a knowledge base in learning theory in order to test and update their own practices. In addition, the acquisition of this language will enable them to articulate their own innate theories about children and learning.

5. Mentors should offer the novice training in the area of planning. They should assist the beginning teacher in articulating the learning target, rather than focusing on content to be covered. This need is also documented in the review of planning studies by Clark and Peterson (1986), and the assessment volumes by Stiggins and
Word Recognition Data

These teachers appear to have grave concerns for a few students in their classes because the students are perceived to have decoding difficulties. The teachers' response is to provide drill and practice, and to test students with flash cards. None of the teachers in this study demonstrates knowledge of a systematic method of transcribing errors for future reflection and instructional revision.

Therefore, it appears that training in running records should be part of initial inservice provided by school districts in transition toward whole language programs. Moreover, on-site support for the use and interpretation of this strategy should be provided. In this way teachers may feel more comfortable during the transition process. They express concern that they will not be able to document the growth of their students in books without controlled vocabulary. Running records would give them a tool to "get a handle on the word recognition thing," as Stacy expresses the concern.

The contribution of these teachers is recognized, however, for they point out that it is not necessary to collect the same data with the same frequency on every child. Fluent readers may only require running records to be performed twice a year, in order to provide data to monitor continuing growth, while struggling readers may need a running record more frequently to match student to text and to provide direct instruction in reading strategies.

Teacher Change

It is not clear from the data gathered in this study, whether teachers' reluctance to make methodological shifts from basal to
whole language practices is due to external mandates in the area of assessment, the fact that their articulated needs for data once their students are in tradebooks have not yet been addressed, or their lack of training in appropriate methods of collecting and using different types of data. Additional studies of assessment practices by teachers in a state of transition are needed. These would provide an opportunity to consider teachers' rate of transition as a function of the pace of changes in external assessment mandates and additions to teachers' knowledge bases.

Finally, school administrators should consider carefully the sequence in which inservice topics are presented to teachers who are making the transition to whole language practices. Assessment must be treated initially or concomitantly with instructional methodology issues (not as an afterthought) in order for teachers to have an integrated knowledge base of curriculum-instruction-assessment.

Most importantly, inservice must first address the underlying belief system regarding literacy. Without this, methodological changes are likely to be superficial accommodations. Experiences with the Open Classroom paradigm of the 1970's taught educators this lesson: meaningful instructional change must be centered around how they think about what they do. If it is not, innovation is merely fashion.

Interactive Data

Elaine succinctly states the dilemma of these subjects: "then there's their daily performance. . .and I don't know how you put that on a piece of paper." It appears that none of these teachers documents interactive and observational data such as oral responses, recreational reading choices, knowledge of vocabulary, cooperative group efforts, or behavior during writing. This is most likely because of time constraints. In addition, they fear such records

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
will be labeled judgmental. Of even greater concern, at least three of the subjects expressed the view that narrative comments on a student could make them libel.

Therefore, it appears that systematic training in selecting data to answer specific instructional questions is needed. Then data collection strategies to capture that information should be provided. These same data could be useful in screenings and eligibilities for gifted and special education programs if it is informative and answers key instructional questions. These teachers know a great deal about their students. What they know should be part of the data mix for high stakes decisions made about their students.

Rather than delineate a list of instructional questions and data collection strategies, it is more appropriate to deal with one or two specific examples. R. is in a fifth grade class. He appears to decode adequately with content textbooks on his grade level. However, he is very distractible and frequently fails to comprehend fiction. This may be due to the fact that stories have gotten longer in fifth grade and his class is working on a novel. The instructional question the teacher poses is "does he comprehend?" In order to find out the answer, we talk with R. We ask him how he likes this book, what he has learned so far, what books he likes better, and anything he knows to do to improve his understanding of text. He is then asked to reread only the first chapter and then asked to write a commercial for that chapter so someone would want to read the book. He is told this purpose before he reads about the assignment. Before reading the next day, he is asked if he would like to read the commercial on Chapter 1 before reading Chapter 2. He is told that he will have to write five questions about Chapter 2 for another student to answer. He must write both the questions and the answers. Each day he is asked to evaluate the written activity.
and to tell if some of these strategies improved his understanding and why.

J. is a girl who is highly verbal. She chatters away like a child several years older. Her sight word vocabulary is very weak. The teacher wants to collect data to understand why her oral language far exceeds her ability to read developmentally. Her invented spellings do not bear any resemblance to the word. For J., assessment takes the form of a running record done once a week. In addition, sound symbol relationships are measured. On running records, J.'s errors are largely substitution of known words with similar graphophonic configurations. In addition, errors in context appear to be small words such as that, what, what, why, when, in, off, etc. On flash cards, J. demonstrates knowledge of all initial sounds. The teacher asks the nurse to check her vision and hearing.

The dilemma of the special educator is a particularly painful one in this transition period. While many of these teachers operate within a holistic, constructivist set of beliefs, they are also bound by legal mandates to evaluate children with instruments that provide quantifiable performance indicators. Sumner (1993) proposes a logical combination of standardized tests, performance testing, observations, student interviews, collection of written products, and anecdotal records to provide eligibility committees with a greater opportunity to understand what children do know as they acquire literacy.

Summary

The results of this study appear to be generalizable in the areas of informing preservice programs and mentorship programs for novices, and informing inservice programs for teachers in transition to whole language. The most crucial recommendations for preservice
teachers are that they should be provided with opportunities to collaborate with master teachers and that they should observe in a school during the first week of the fall semester. In this way they will be able to gain knowledge about managerial decisions related to organizing a classroom for the year.

The results support the studies of novice teachers (Hollingsworth, 1989; and Doyle, 1979) and planning studies (Yinger, 1980). Novices attend to managerial decisions during their first year(s) and rely on paper-pencil data for assessment. They become aware of interactive data only after classroom management becomes routine. Recommendations for mentorship programs for novices place the mentor in the new teacher's classroom. This is the context in which the novice is most likely to acquire and implement new knowledge in daily practices.

Teachers in this study appear quite typical as older, experienced professionals who have not been early adopters of whole language. This supports the survey work of Barry (1992) and Coulter (1992). These results also reveal that many experienced professionals have supplemented a basal program with whole language strategies. In order to make the transition to whole language, rather than merely assimilating whole language strategies within a basal program, it is necessary for teachers to acquire a knowledge base in authentic assessment and in reading as a constructivist process. They must learn to differentiate between data for instructional monitoring and data for accountability. A reading curriculum is needed with specified learning targets. Groups of professionals should determine what data provide clear, convincing evidence of mastery of these targets. Then teachers must be trained to collect, document, value, and report these data.

Moreover, these teachers appear preoccupied with accountability concerns. Administrators and teachers must dialogue with parents.
during this transition phase in order that the latter understand that descriptive data will tell them more about their child's literacy development. Grades must reflect progress toward learning targets without the "noise" of noncognitive factors. Reporting systems must provide separate information about academic performance from noncognitive factors and provide information comparing the student's progress to the learning target as well as about the student's relative standing in the class. Finally, teachers must be allowed time and support within the school site in order to confront their belief systems about literacy and assessment. In this way they can truly integrate new knowledge and skills into their daily classroom practices (Borko, in press). Borko's work on teacher change also suggests that this will take longer than a school year to accomplish.
References


teachers predict the scores of their pupils in standardized tests of attainment? A study of some non-cognitive factors that influence specific judgments. *Educational Studies*, 11 (1), 41-60.


Clearing House. 61(7), 320-324.


Appendix A

Interview Guide
Below is an outline of the open-ended interview sessions. Please note that data is not specified (i.e., "do you value students' oral responses in class", etc.) to avoid leading the informants.

Session #1

State purpose of study. (The questions you will be asked will be related to how you gather information in your classroom on your students. There are no right or wrong answers. Different teachers use different information in making their decisions. The purpose is to find out what information you use and how you modify instruction based on this information. This will assist us in training new teachers and in developing new assessments in the area of reading. This is truly an effort to gain an understanding of how you think. Please be as open and as honest as possible. The information you provide will not be reported in an identifiable fashion nor will it be used in any way to evaluate you. I will be happy at any time to provide you with a transcription of the notes from any of our
sessions. I will return to you any tapes made at the conclusion of our study.

Ask informants to save any notes they make to themselves regarding assessments or documents that illustrate thoughts they share during the sessions.

Please describe your reading program. (Follow up questions may relate to use of basal, literature, writing, etc. to gain an understanding of this teacher's orientation and philosophy of reading as well as building/County mandates which may or may not coincide with his/her philosophy.)

Please share and explain any groupings you have in your classroom in reading. How were these formed? How do they change? Who has changed groups? Please explain how this decision was made? Are there any students who may change groups in the future? Why? (With all further questions, follow up questions will be designed to elicit sources of data which confirm the teacher's attribution for these students. The method employed will be to focus upon the
students in all but the last session in order to encourage the teacher to "think aloud").

How are your students progressing in reading? Who is not progressing in their present group? When did you first notice this? What is the greatest need for this student? What is their area of strength?

Session #2:

Please think back to the student you described last time who was having difficulty in reading. (Read back transcription or notes). Could you show me some of this student's work and talk some more about him/her. What can you usually tell from a student's work?

Could you relate some instances about this student's classroom learning behavior in the last few days that illustrate what you are saying?
Is there any other information you have on this student? How does this help you in teaching him/her?

What is your diagnosis of this student's difficulties? What do you think would help?

What other information would you like to have on this student if you had time to spend with him/her individually?

Set up session #3: Before our next session, we will arrange to have a 10/20 minute segment of a reading instructional activity videotaped. Please select a session that you think is a good opportunity for you to assess your students. The video taping should focus upon the very best readers in your class. We will discuss them next time.

Session#3:

View tape. As we view the tape, please stop it each time you feel you have made any assessments as a teacher.
(Probe specific interactions that are not discussed by the teacher.)

Are these students equally good readers? Why or why not? Who is the strongest? Why? Who has more needs? Why.

Set up #4: Please pull five cumulative folders, reading folders and work sample folders (or portfolios) if you keep these.

Session #4:

Please leaf through a student's reading folder and his/her accompanying folders and talk about the needs of this student as you seen him/her.

Do these records accurately describe the student as you see him/her? Why or why not?

Set up #5: Next time we'll look at your grade book and talk about how you assign grades.
Session #5:

Please share your grade book. How do you grade in reading? What do the grades tell you? Who is making the poorest grades? Why?

Please share the grades for the "At Risk" student identified in session #1 and the strongest students in the video. (Encourage the teacher to talk about their grades and the information she derives. Attempt to probe what mix of cognitive and non-cognitive measures are reflected in the grades.)

What do you think would improve these students' grades?

Session #6:

Pre-Observation interview: Please explain what I will be observing. What are the special needs of these students. How do you predict they will do with the lesson you will present? What makes you think that?
Session #7:

Observation

Session #8:

Post-observation interview: Review data from #6. Ask teacher to reflect on lesson. How did specific students perform? How do you know? What data support this?

Were there any surprises during the lesson? Why?

(Probe whether notes are made during such episodes or how judgements made during instructional activities are documented.)

Session #9:

What conscious strategies do you use to assess student progress in reading?

Who influenced you the most in developing these strategies?

How have your assessment strategies changed from the time when you first began teaching?

What sources of data do you value less than you did as a beginning teacher? Value more?
What have you found over the years is not a reliable source of data on your students?

As you spend more time with your students as the year progresses, what sorts of things do you notice? Why do you think you didn't notice this before?

What procedures do you use in September to "size up" your group? How have your initial estimates changed for this group of students? Why? (Probe factors relating to ability, prior achievement and common sources of bias from the literature.)

If you had time, what information would you like to know about students? How would you collect this?

If the father of one of your students who lives across the country came to visit your classroom after school, what would you show him that would enable him to understand what you want second graders to know and what his child can and cannot do.

What information would you collect to strengthen your judgement about a student if you were being challenged by the parent? (What pieces of data give an accurate snapshot of a child at a point in time?)
Please share what information the school division and/or building level administrator requires you to collect? How is this information used? How would you change this collection or documentation of information?

What do you think the beginning teacher needs to know in order to accurately assess reading achievement? How do you think this could be accomplished? What advice would you given them? If you were assisting with inservice for these teachers, what would you include in the presentation?

Session #10

Share a profile of decision making strategies and the rank order list of assessment data utilized.

Ask for feedback.

Explain further analysis of this data and reporting format.
Appendix B

Examples of Interview Data
(Could you describe your reading program...the groups you have... the materials you use... how you're organized... scheduling... that kind of thing.)

using HBJ materials... the second grades at our schools for the first time is grouping for reading... so I have the top students... I have 31 students... which is a lot of boys and girls... but they're very motivated...

there are 2 students who are in celebrations (3rd grade beginning text) which is level 8 and the other two groups are in Windmills, 7-2 and
Exemplar:

are they tested on them (words on basal units) any more? no

Do they need to be tested on them anymore? I don't know... all I know is I want my children to be able to know what the word means, use it in a sentence correctly, and be able to read it...

(orally or in writing)
use in sentence orally..

Tag(s): word recognition/decoding/word meanings/oral language/ability to use word in sentence orally

Exemplar:

and know when you see it what it means... so that when you get to another word... and you can't figure it out... you know the context of the sentence... and you're able to go from there....

Tag(s): ability to use context clues to facilitate word recognition

Exemplar:

so on my vocabulary sheets that I send home, it says, "please make sure that your child can read this word and use it correctly in a sentence."

Tag(s): word recognition/ability to use word in sentence/oral language/word meanings
uh-- I take about 1 or 2 a week

and so I have a lot of grades....because I want to give the benefit of the doubt... I don't just give 'em 5 grades ... cause that's all you have to have...I give them a lot... and sometimes it's reflecting... like this was a whole week.. this was actually 6 or 7 days in here... the Freckle Juice... and so...

Tag(s): grades/pattern of grades/number of grades/reliability of grades as indicator

Card I.D.: 26024

Interview No.:   Ques. No.:   Source Card I.D.: 2512
Source: stack

Exemplar:
many grades are important to her for reliability-- "I want to give the benefit of the doubt"

values grades more that reflect greater period of time-- such as Freckle Juice cumulative grade

Tag(s): grades/reliability of grades/grading over time/cumulative vs. daily grades

Card I.D.: 26223

Interview No.:   Ques. No.:   Source Card I.D.: 8312
Source: stack

Exemplar:
Observation: Children were working on an HBJ vocabulary chart in a small group at the back table. The rest of the class was working very quietly at their seats.

The first word was "hero". She asked for meanings of hero. After one child said, "help a person in trouble", she wanted more and restated the child's response.

Tag(s): oral response/word meanings/vocabulary

Card I.D.: 26417
Note: prior to this interview I provided the informant with a written transcript of previous interviews. At the beginning of this interview she inquired about the use of the word "judgement" in relation to recreational reading choices of her students. She wondered whether this was a positive or negative statement. I assured her it was neither and that my intent was merely to point out that she was sensitive to this data— the choices her students made in the classroom for recreational reading. I felt this was significant. I truly doubt many teachers attend to what their students are reading because they are attending to an instructional group. I felt the statement showed that this informant has mastered the art of "simultaneity"— e.g., she can attend to simultaneous pieces of data floating around in her classroom. (see Hollingsworth on beginning teachers)
let's open our books now to page 134...
(this was the end of the video tape)

(it's interesting that you chose the vocabulary... you feel like that when you hear them define words... that really gives you a lot of insight?)

uh-huh--

(into their reading skills?)

yes

(or more into their comprehension skills?)

well, I think it ties all in... you know... oral language... to get them to express themselves... cause if they can't express themselves... then... I think they have difficulty understanding what the printed word is.... because they really wouldn't know the context that it is saying...

(what about comprehension?)

it's not really the best... it's very slow... he's very slow to answer... which I give him a lot of wait time... I just sit there and wait.... and he'll usually say something... but it's very difficult for him...

he's ADHD ... he's on medication

it's helped a lot... from last year... they had a lot of problems

I think they're improving... I really see them improving... where they couldn't read at all when they came to second grade...
April 1993 Observation Notes

Each child had a book. The teacher read - the children followed.
"Get your partners" They got into pairs, read aloud to each other.

The teacher chose 3 children to read aloud with her - Why? The oral reading of quality of their work - different.
2 children worked in the back and didn't participate - Why? - (bee) Their needs are different - They are in "Super Six!"

Read responses - Chapter by Chapter.
Reading Log -

10:15 p.m.

Although this was Whole Group, modifications were made for 9 students.
Observation 5-27-93

- HBJ unit b-4 Vocab + comp only -
- 10 children
- provided an indep. activity when testing
  is over
- "if you have a problem I'm
  right here"
- "WTH" it was city-
  clue: "What is Wunsbg?"

- ¿ 7 country repeated

- children came for w.r. assistance
- each time the teacher gave clues -
  usually close clues
- ex: celebrate -
  "What do you do on your birthday?"
  (calendar)
  "No-- there's a calendar up on all yr.
  you invite people over to help you"

- "finally came & sat by
  the teacher -
  she appeared
  eager to perform correctly -

- "I forgot WTH is "
  "cover the ed"
  "dix"
  "that's the way it begins"
  "What is the vowel?"
  "ok? WTH?"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
2-18-93

Interview Notes

T

Compared Sept. writing & writing now -
length of story -
depth of story -
new story has middle --

Journal Entry

L - dramatic change -
was level 3
had made good progress in reading - her writing reflects this.

R

has moved down a group -- failed reading test
no home support --

A

As they gain reading proficiency spelling & punctuation is more adequate.
exit interview
6.24.93

as new teacher: lot of weight in level test:
we tested skills more
I've learned how to make modifications
to meet test taking needs
distractibility
read orally
strategies to read & find answer--
"I was much more rigid 1st year."
"I never read words for them."

oral reading & how they do in groups.
new teacher: tremendous amt. of
what work
monitor work habits--
kids would sit & complete work
# Reliable tool--

beginning -- did not have inf. about
IRT.
always felt environment & experiences
were important--
their own language
if their speaking vocab is limited,
then their reading

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
12/7/92 Betty Field Notes

talked about themes of:
- grading to provide student motivation
- grading to respond to parent's expectations
- pressure of levels on students

Relate these issues to the "noise of assessment". These non-cognitive influences on valuing data may bias grading -- are these distinct from assessment?
he doesn't understand the concept of his daily routine..."
See p.19

[Now she describes a language processing prof. & workers about which came first -- lack of underst. of his environment or behav. problems.]

"so it's almost a defeated attitude." describes frustrational behavior of being unable to read with his creative story.

52+53
She observed Ch. sit & do no work. She concludes that their seatwork is less valuable than a tool than oral responses in grp.

p.1 (non-cog)
Child acts out related to custody battle.
Appendix C

Examples of Notes and Documents Provided
A

writing

Original sentences with vocabulary words

I am tall.

I see everything.

I hit a Ball.

The boy has Ears.

I build a home.

I have a head Roc.

Their boy has shoulders.

I bet you cannot run.

You cannot fly.

You crashed into a car.

I am short.

Jan. 17, 1993

Scotty

need more in sentences.

looks for meaning

looks for grammar

length of sentences

reveals oral language.
Close work sample

The boy jumped off the house.

They are jumping very high

Who will jump next?

We stopped working.

Will you stop talking.

Standing at the door.

\( \frac{4}{10} \)
Creative Writing

The dog and the cat

Feb 18, 1973

One day a dog was chasing a cat. The cat was tired of the dog and the dog was tired of the cat. One night the cat wished for him to be a dog for 100 hours. He wanted the dog to be a cat for 100 hours. The dog got a taste of his own medicine. The cat chased the dog all the time.
Early Journal Entry
Sept. 8, 1992

My Summer Week to the pool
Abraham Lincoln

Abraham was a president.
Abraham got killed by a gun named John Wilkes Booth.
Abraham lived in a log cabin.
Abraham had a beard because he did not shave.
Abraham loved to read.
He would set by the fireplace and read.
Date

Dear Needles Park

I love a passed out paper

Sincerely,

[Signature]

1-6-93
TO ____________________________ DATE ____________

✓ GOOD

- NEEDS IMPROVEMENT

____ Motivation
____ Daily Work
____ Homework
____ Self-control
____ Follows directions
____ Other

____ Attitude
____ Participation
____ Attention
____ Cooperation

SUBJECTS/SKILLS

______ Reading
_______ Language
_______ Math
_______ Spelling
_______ Social Studies
_______ Science

____ Conference requested ____________ Date ____________ Time ____________

________________________________________

Teacher

Please sign and return

Parent

Communication

NOTES HOME

sent home every 2 weeks 3-31-93
Autobiographical Statement

Margaret Mary Dinan Davis

B.A., June, 1964, College of William and Mary
M.A., August, 1975, College of William and Mary
C.A.S., May, 1987, Old Dominion University

Margaret Mary Dinan Davis was born on July 5, 1946, in Richmond, Virginia. She is an educator with twenty-seven years of experience: seven years as a first grade teacher; sixteen years as a reading specialist; and four years as an administrator. All of her experience has been in the York and Williamsburg-James City County School Systems in Virginia.

She is currently employed in the York County School System in Virginia where she is an assistant principal at Coventry Elementary School. She has conducted numerous presentations at local, state, and national conferences on topics of emergent literacy, assessment, pacing of instruction, volunteer tutoring programs, creative dramatics, clowning, and reading research. Her publications include articles on creative dramatics, the hammered dulcimer, and student travel abroad. In addition, she has published several original musicals for children.

She is a liturgical musician and performs professionally throughout the state as a member of the duo, Irish Aire. She plays hammered dulcimer and many other instruments and performs a repertoire of Irish and Appalachian music. She has recorded three albums, Earthtones, Rosebower, and Wintercarols.

She is the recipient of a competitive research grant award from the Virginia Educational Research Association and grants from the Virginia Commission on the Arts. She co-authored a paper presented at the annual conference of the American Educational Research Association in 1990. She was the 1991 Reading Teacher of the Year for the Newport News, Virginia Reading Association. She is listed in Who's Who in American Education, 1994. She is a charter member of the Gamma Phi Chapter of Delta Kappa Gamma and currently serves as president of the chapter. In addition, she is a member of Phi Kappa Phi honorary society.