



This Brief is based on the following CTP Occasional Paper:

Meeting the Needs of Failing Readers: Cautions and Considerations for State Policy

April 2003

Marsha Riddle Buly
Western Washington University

Sheila Valencia
University of Washington

This document can be downloaded at the CTP web site, <http://www.ctpweb.org>

Another version of this research document, *Below the Bar: Profiles of Students Who Fail State Reading Assessments* was published by **Educational Evaluation and Policy Analysis**, Fall 2002.

Understanding Reading Test Failure: Challenges for State and District Policy

Every year thousands of students fail state reading tests and every year policymakers and educators search for strategies to help these students succeed. Discussions about possible interventions or “solutions” to this problem are often driven by test scores, and the policies that emanate from them often recommend specific programs or strategies for all failing students. Interestingly, up until now, the tests that have launched these conversations typically have been administered at the 4th grade or higher—a common benchmark year for both national and state assessments. Tests such as the National Assessment of Educational Progress (NAEP), state developed assessments (e.g., ISAT or FCAT), or state mandated norm referenced tests administered at 4th grade (e.g., SAT-9, ITBS) target reading comprehension. Nevertheless, many of the suggestions for intervention focus on decoding skills or phonics. The assumption seems to be that students who fail reading tests at 4th grade and above lack more basic, earlier reading skills. Yet, little empirical evidence supports such an assumption.

One State’s Response to Reading Test Failure

The situation in Washington State is similar to that in many other states. Although student performance in reading has improved somewhat over the past four years, too many students continue to fall below state standards. In 1998, the second year the new Washington Assessment of Student Learning (WASL) was administered, 43% of 4th graders failed to reach proficiency. In response to the test scores and to the outpouring of research on early reading instruction and intervention, the legislature provided extra funds to support beginning reading instruction in K-2 classrooms. While the legislation called for a comprehensive approach to instruction, some public officials and school districts endorsed a focus on phonics, a stance that was consistent with the emphasis in many other states at the time. The heightened interest in phonics and in other specific instructional programs and approaches for struggling readers has resulted in a proliferation of both informal and formal school district policies on reading instruction.



The situation and response to it begs two critical questions: 1) Does the “solution” fit the problem? 2) What precisely is the problem? To find out why students fell below standards on the WASL in reading, more than 100 4th grade students from a diverse semi-urban district participated in a series of individual diagnostic reading assessments. None of the students had reached proficiency on the state assessment and none of the students was currently receiving supplemental instruction in either Special Education reading or English as a Second Language (ESL). The sample did include students who had been exited from ESL programs or whose family language was not English but who did not qualify for additional support.

A team of seven educators administered five reading assessments to each student which provided multiple indicators of students’ abilities on several different aspects of reading (e.g., reading accuracy, expression, rate, comprehension, etc.). In general, these tests provided information on five major components of reading: word identification, phonemic awareness, comprehension, reading fluency, and vocabulary. In addition, information about students’ writing ability, home language, and socio-economic status was also collected. Students took the state test in April of 4th grade and all diagnostic assessments were conducted in September-October of 5th grade, immediately after state test results were available.

Behind State Test Scores: Profile of the Group

The results of all the diagnostic tests given to all the students were analyzed to determine if they could be combined into a smaller set of factors to account for students’ failing performance. Three factors—meaning (comprehension and vocabulary), word identification, and fluency (rate and expression)—accounted for the great majority (78%) of the variance on the WASL scale scores. Average scores for the entire group of students revealed that students’ word identification skills were beyond a basic, beginning level. They scored at approximately

4th grade level when reading words in isolation and at the 3rd grade level when reading words in passages. In the area of meaning, students averaged 2nd grade level in comprehension of reading passages and scored somewhat below established age norms for vocabulary. Average scores in fluency, both rate and expression, were substantially below expected levels for 4th grade students.

The average scores suggest a particular pattern of student performance—one in which students’ performance is slightly below grade level in word identification and substantially weaker in comprehension and fluency. However, average scores obscure important individual student differences.

Behind Group Averages: Profiles of Individual Students

Students may be failing to meet standard on a state test for different reasons, and consequently, they may require different instructional strategies. By using cluster analysis¹, the study determined that the pattern suggested by the averages for the whole group did not fit all, or even a majority of the students. Rather, six distinct student profiles emerged demonstrating that students were actually strong in some areas and weak in others. Each profile is described below, drawing on actual data from prototypical students.

- **Profile 1: Automatic Word Callers (18%)**

The students in this group read the words quickly and accurately but failed to read for meaning. More than 60% of these “automatic word callers” were English language learners and most were from low-income families. For example, Tim, a native English speaker, was very good at identifying words and applying phonics to nonsense words, scoring on a 9th-grade average on word identification tests. His reading rate of 174 words per minute was among the highest of the students in the study, yet his comprehension scores were far below grade level. Another student, Marja, had

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R308B970003, as administered by the National Institute on Educational Governance, Finance, Policymaking and Management, Office of Educational Research and Improvement (OERI), U.S. Department of Education. The contents do not necessarily represent the positions or policies of the National Institute, OERI, the U.S. Department of Education, or the endorsement of the federal government.

similar strengths in word identification and fluency, as well as weaknesses in meaning. However, she had difficulty with meaning primarily because she was still developing word meanings (vocabulary) in English; like many other students in this profile group, Marja was a second language learner. Although both of these prototypical students fit the profile of automatic word callers, each would need a different instructional approach.

- **Profile 2: Struggling Word Callers (15%)**

Although similar to Automatic Word Callers in terms of demographics, strong fluency, and difficulty in comprehension, the Struggling Word Callers had considerable difficulty with word identification. One student, Dang, whose home language was Vietnamese, had uneven word attack and word identification skills; he was better at reading words in lists than reading words in the context of longer passages. He read at a relatively fast rate, apparently believing that good reading is fast reading, yet he failed to self-monitor or self-correct when his reading didn't make sense. This difficulty with comprehension was probably compounded by Dang's difficulty with English, as his low vocabulary scores suggested. He would likely need instruction in word identification, oral language, and comprehension, especially metacognitive strategies.

- **Profile 3: Word Stumblers (17%)**

Primarily from English speaking families, the students in this group were slow readers who also exhibited considerable difficulty with word identification. Sandy, for example, scored at the 1st and 2nd grade level on measures of word identification. Although this would likely interfere with comprehension, Sandy spontaneously self-corrected many of her errors when reading passages, trying to make sense of the text. As a result, Sandy had strong comprehension but her fluency scores were low. She would need instruction targeted at word identification as well as fluency building through reading text at her instructional level.

- **Profile 4: Slow and Steady Comprehenders (24%)**

These students read slowly even though their word identification and comprehension abilities were relatively strong. For example, although Steven's word identification scores were close to grade level, he read texts slowly, appearing to lack automaticity in decoding and to have some difficulty with reading more complex multisyllabic words. Also in this group was Stacey, who was stronger in word identification than Steven but whose reading rate was also substantially below standards. Although both these students have strong comprehension, their slow rate is likely to limit the pleasure they get from reading or the time they spend reading. As reading demands increase in school, these students may fall further behind. Both would need to engage in more reading of relatively easy material to build their automatic word identification and rate of reading.

- **Profile 5: Slow Word Callers (17%)**

These students were word callers like those in Profile 1 but they were not automatic. They were accurate readers who were slow and who also struggled with meaning. From the existing data it is difficult to determine just what is contributing to student performance in this profile. On the one hand, students who read very slowly may experience comprehension difficulties. On the other hand, students who recognize that they are not understanding should slow down to try to make sense of the text. Some students in this profile, like Joey, demonstrated difficulty with comprehension, vocabulary, and fluency. Others, like Jose, demonstrated difficulty with only comprehension and fluency. Both students would need more diagnosis to disentangle the source of their comprehension and fluency difficulties.

- **Profile 6: Disabled Readers (9%)**

Students in this group were low in all three areas—word identification, fluency, and meaning—but they were exceptionally low in word identification, scoring at the 1st grade level

and below. With such word identification skills, fluency and comprehension difficulties are likely to follow, creating students who are struggling on all fronts. These students would need interventions beyond what most classroom teachers could provide.

Policy Implications

This study confirms that students fail reading tests for many reasons. Each student has a distinctive pattern of performance that can help teachers target reading instruction and, consequently, improve a student's reading ability. The results reported here provide direction for policies—at state, district and school levels—that will likely improve student achievement rather than merely measure it. The findings suggest five potential “policy levers.”

The **first policy lever** is support for differentiated reading instruction based on in-depth analyses of reading test performance. This study raises questions about mandating specific instructional strategies or programs, such as phonics or fluency building, for all failing students. In this sample, for example, phonics instruction would have been inappropriate for more than 50% of the students, those whose problems were comprehension, fluency, or language rather than word identification. Similarly, an instructional strategy that did not address fluency would have missed the needs of almost 70% of the students.

Instead of mandating particular types of instruction based on average group scores, policies should help districts, schools, and teachers undertake deeper analyses of students' abilities and exploration of a range of instructional strategies that could meet students' needs. For example, school districts and schools might be required to engage in, and report on, a self-study of student reading performance as part of their improvement plans. The study might ask such questions as: Which students did we predict would do well on the test and did not? Why did they fail? What evidence do we have about the progress of our lowest performing students? Which instructional practices are we using that seem to be most effective? What other measures of students' reading abilities do we have and what do they tell us? A policy that supported a self-study around test scores such as

this would require more data collection and finer-grained analyses of the problem and possible interventions.

A **second policy lever** is the effective use of multiple indicators of reading ability. No state assessment can provide the detailed information needed to understand complex reading performance nor can it provide information quickly enough to inform daily instruction. In short, there needs to be greater reliance on classroom-based assessments. Some states such as Vermont, Kansas, and Washington have implemented classroom-based assessments at early grade levels and have supported teachers in learning how to administer the assessments and interpret the results. These assessments are diagnostic; tailored to individual student needs; individually administered, scored, and analyzed by classroom teachers. Simply mandating such assessments, however, is not enough to improve teaching and learning. Not only must teachers have the knowledge, support, and authority to use the assessments well, but they must also see the assessments as helpful to their instruction. Policies that support classroom assessments and the professional development necessary to use them effectively would be useful at all grade levels.

A **third policy lever** is flexible alignment among standards, assessment, and instruction. Alignment is considered necessary to standards-based reform, but sometimes it can be oversimplified. Assuring, for example, that 4th grade teachers are teaching the 4th grade content standards does not necessarily mean that they are providing appropriate instruction for all students. For some struggling students, teachers may need to draw from lower grade level standards to match students' needs and use actual grade level standards as goals to guide their teaching. State and district policies should provide curriculum frameworks to help teachers articulate both grade level expectations and developmental perspectives on teaching. The policy and its message should be explicit—the aim is not simply to teach the standards, it is to teach students with different learning needs so they may eventually reach grade level standards.

The **fourth policy lever** is local control over allocation of resources for curriculum materials,

time, instructional interventions, and professional development. If teachers are to tailor instruction for a wide range of student reading abilities, they need appropriate materials (including several grade spans for such subjects as science and social studies). Students reading below grade level need appropriate instruction and materials throughout the school day as well as more instructional time and more targeted instruction than other students. In addition, the results of this study raise concerns about the instructional support provided for English Language Learners. Although the study did not include students who were receiving supplemental ESL instruction, 43% of the failing students in the sample were English Language Learners who had been exited from the support program. Clearly the resources allocated for ESL instruction need to be reconsidered. Overall, policy that supports sustained, on-site, content-specific professional development is critical as teachers develop understandings of reading processes, reading instruction, and teaching in complex settings.

A **final policy lever** is support for evaluations that reflect the complexity of growth in reading proficiency. Many state tests are not sensitive to growth of students who are reading far below their grade placement or to changes in teaching practice that may not have immediate impact on student learning. Policies need to support the development of a system of intermediate indicators of student learning that will demonstrate student growth,

especially for students who are not yet meeting grade level standards. The data will need to draw from greater use of diagnostic classroom assessments.

In sum

A single, group-administered state reading test can not provide the kind of detailed information needed to understand reading failure. Similarly, no single intervention can possibly address the wide range of needs of struggling students. If standards-based reforms and the investments in them are to improve learning rather than simply measure it, then policies must support building the capacity of educators to identify and understand reading performance, and to respond appropriately and flexibly to meet students' needs.

Endnote

- ¹ Cluster analysis is a statistical procedure in which the most similar pair of cases is initially combined, and each new case (individual student) that best fits a cluster is added until all are placed. This study averaged each student's standardized variable scores for word identification, meaning, and fluency and used those scores in the cluster analysis.



Center for the Study of Teaching and Policy

A National Research Consortium

University of Washington (lead institution) • Stanford University • University of Michigan • University of Pennsylvania • Teachers College/Columbia University

Other active participants in CTP's research and dissemination program include researchers affiliated with Indiana University, Michigan State University, Pennsylvania State University, the University of California at Santa Barbara, the University of North Carolina, and Education Matters, Inc.

CTP studies the way policies and conditions in schools, districts, states, and the nation shape the quality of teaching and learning in our nation's elementary and secondary schools. The Center pays particular attention to the ways these policies and conditions interact with each other to influence the teaching profession and its practice.

The Center's program of research is carried out in collaboration with various other research organizations, among them other OERI-funded research centers, including the Consortium for Policy Research in Education (CPRE), the Center for Research on Education, Diversity, and Excellence (CREDE), and the Center on English Learning and Achievement (CELA). CTP is affiliated with a variety of professional and advocacy organizations which represent teachers, teacher educators, state and local policymakers, disciplinary groups, and educational reform interests.

CTP Teaching Quality Policy Briefs

One of the Center's chief goals is to offer timely and useful information concerning efforts to improve the quality of teaching to national, state, and local policy communities. The Teaching Quality Policy Brief series is one vehicle for accomplishing this purpose. Issued periodically throughout the year, these briefs summarize Center working papers, reports, occasional papers, and other research products. The original publications, as well as the briefs, are available for download from the Center's website: www.ctpweb.org

CTP Policy Briefs:

- #1 **State Action to Improve Teaching**, December 1999
- #2 **State Teaching Policies and Student Achievement**, December 1999
- #3 **A Different Approach to Solving the Teacher Shortage Problem**, January 2001
- #4 **Connecticut's Story: A Model of State Teaching Policy**, June 2001
- #5 **San Diego City Schools: Comprehensive Reform Strategies at Work**, February 2002
- #6 **What School Districts Spend on Professional Development**, November 2002
- #7 **The Organization of Schools as an Overlooked Source of Underqualified Teaching**, December 2002

Recent CTP Research Reports and Papers:

- Triage or Tapestry? Teacher Unions' Work Toward Improving Teacher Quality in an Era of Systemic Reform**, June 2003
- Theorizing About Teachers' Responses to Reform: The Role of Communities of Practice**, May 2003
- Leading for Learning: Reflective Tools for School and District Leaders**, February 2003

Contact Information

Michael S. Knapp, Center Director
Miller M201, College of Education
University of Washington 353600
Seattle, WA 98195-3600
Email: mknapp@u.washington.edu

Michele C. Ferguson, Center Manager
Miller 203C, College of Education
University of Washington 353600
Seattle, WA 98195-3600
Phone: (206) 221-41145
Fax: (206) 616-8158
Email: ctpmail@u.washington.edu

Sally Brown, Communications Director
Miller 404B, College of Education
University of Washington 353600
Seattle, WA 98195-3600
Phone: (206) 543-5319
Fax: (206) 616-8158
Email: salbrown@u.washington.edu

Web Address
<http://www.ctpweb.org>



Center for the Study of Teaching and Policy
University of Washington Box 353600
Seattle, WA 98195-3600

Nonprofit Organization
US POSTAGE PAID
Seattle, WA
Permit No. 62