

APPENDIX C

Observations: What Research Tells Us . . .

The material in this Appendix includes references to research as well as research summaries. The major topics are listed below.

Teachers suffer from initiative fatigue; we need to focus:

1. Mike Schmoker, author of *Focus: Elevating the Essentials to Radically Improve Student Learning* (2011), speaks to the need for simplicity, clarity and focus.
2. Michael Fullan, author of *All Systems Go*, (2010), addresses the need to focus at the state level on literacy, numeracy and critical thinking.

Top-down policies may have unintended negative consequences:

3. Linda Darling-Hammond's study on teacher quality and student achievement asserts that the states with fewer policies have higher student achievement, positing that top-down management does not advance teaching and learning. Her full document can be found at the University of Washington's Center for the Study of Teaching and Policy, December 1999. The full title is, "Teacher Quality and Student Achievement: A review of state policy evidence."
4. In an article called, "Cost doesn't spell success for Colorado schools using consultants to improve..." the author demonstrated how ARRA funds were paid to consultants and schools weren't "fixed." http://www.denverpost.com/investigations/ei_19997418.

Standardized assessments alone cannot evaluate expert teaching:

Though capable of being useful, the state standardized test, alone, cannot determine teacher quality and should not be used as the sole measure of teacher accountability. A more useful assessment for teachers day to day is the formative assessment. This assessment can be administered in a classroom, a grade level or a building to use as part of a teacher evaluation system.

5. In Popham's book, *Evaluating America's Teachers, Mission Possible?* (2013), mentioned in the body of this report, explains how to use a valid and reliable teacher evaluation system that improves teaching and learning. The system is essentially a local control issue.
6. Popham, W.J. & Stiggins, R. (2013). *Today's teacher evaluation: A cautionary note*. In this article, the authors caution against the use of standardized test to prove teacher effectiveness. The full article is in a research appendix in this collection.

7. Linda Darling-Hammond's work, "Evaluating Teacher Evaluation," February 29, 2012, which originally appeared in the *Phi Delta Kappan* and *Education Week*, cautions against the use of standardized tests to measure teacher effectiveness.

Criterion-referenced state tests are more instructionally supportive than standardized ones.

8. Jim Popham and his colleagues wrote about the instructional support of the PAWS test in the March 2012 issue of the *American School board Journal*.

Beware of one-size-fits-all federal and state initiatives that can have negative effects on expert teaching. One example is unintended negative consequences of Reading First.

9. Overview by Sheryl Lain of the SEA document form 626 along with a copy of the report to demonstrate the proliferation of commercial programs producing little of success.
10. This appendix includes a summary of Richard Allington's meta-analysis on what works in reading, especially with struggling readers, followed by a chart from What Works Clearinghouse, mentioned in his book. The chart shows that Reading Recovery works. Reading Recovery is a teacher training program not a textbook scripted program teachers use with students.

Two State-level Professional Development Initiatives That Enhanced
Student Learning
By Sheryl Lain

To improve student learning, the professional development delivered in the state can be very effective. During the past year and a half, the effort at the WDE was to alter the delivery of instructional support to schools in two ways--a statewide approach and a school-based approach.

3+8: Statewide Training on Neglected Portion of PAWS Reading

In 2011, Wyoming used a peer-to-peer delivery system and the results are notable. When teachers across the state took part in a statewide professional development in reading, reading scores on the state assessment improved 12% in two years, evidence that when teachers work for common goals, their energy to teach is supported and student performance improves. Other data confirms that teacher-to-teacher delivery systems produce positive results. In 2011, over 60 schools did not meet Adequate Yearly Progress in special education; in 2012, the number dropped to about 40 schools. One school with a School Improvement Grant (SIG) saw astonishing reading growth in the third grade cohort from 18% proficient in 2011 to 58% proficient in 2012 (*Billings Gazette*, 2012).

What happened?

Using a state-level professional development delivery system, a cadre of expert teachers, those with evidence of student success and a background of facilitating professional development, were contracted to hold weekend trainings for over a year in local schools in every geographic area of the state. The trainings were taken right to the teachers' back yards. Participants extended the weekend work back at school through sharing blogs, writing and posting responses to research, and viewing and responding to teaching videos. About one-fifth of Wyoming teachers attended and the evaluations were positive with a 4.79 approval rating out of 5 points.

Both the high evaluations and the subsequent rise in state reading scores demonstrate that teachers learn best from their excellent, experienced colleagues who listen, share, model and coach, passing the torch from one to the other across the hallways and across the state.

What does a review of the literature say?

The teacher-to-teacher method of delivering professional development is not new. A literature search reveals that peer-to-peer delivery underpins the successful National Writing Project, a premier professional development system developed decades ago with clear evidence of success. The National Writing Project (NWP) is the oldest professional development model in America begun in 1973 and funded by Congress for almost three decades. The Council for Basic Education stated that the NWP is the most successful and far-reaching initiative to improve literacy in America. Quoting from their backgrounder, "The NWP operates on a teachers teaching teachers model" (www.nwp.org). The teachers who served as facilitators for the state delivery system were almost all former Teacher Consultants for the Wyoming Writing Project, affiliated with the NWP.

Another initiative based on the peer-to-peer delivery system was rolled out in 2004 by the U.S. Department of Education (USDOE). Using the moniker *T2T*, the USDOE delivered best practices to 20,000 teachers in 50 American cities (www.paec.org/teacher2teacher).

Other highly successful peer-to-peer delivery systems include Reading Recovery (www.readingrecovery.org) which makes use of expert practitioners coaching their peers. Even teacher evaluation systems such as Charlotte Danielson's (2007) utilizes teacher mentoring as a proven method of developing teachers. Finally, across the country, successful schools are making use of formalized teaching communities wherein grade level teachers routinely congregate to conduct book studies, examine student data, share student products and set common goals (www.centerforcsri.org). Again this is an example of peer-to-peer or teacher to teacher delivery systems at work.

What are four advantages to this state-level delivery system?

A peer-to-peer delivery system for professional development has four advantages. The method is sustainable, consistent, practical and fiscally responsible:

1. This method is sustainable because the teachers who teach and learn return to their schools and not only practice the newly acquired knowledge in their classrooms, but also share with their peers in learning communities. Attendance at the trainings is voluntary and held on weekends to avoid pulling teachers away from their classrooms. For certification and graduate credit, teachers extend their learning beyond the ten-hour weekend, the better to embed the content and pedagogy into routine practice. Extensions include working in school-based learning communities watching videos together of expert teachers in classrooms; reading research and discussing with other educators by blogging; selecting, using and sharing lessons posted on the web; and taking further trainings to deepen practice. In this way the cycle of continuous school improvement is fostered--all with the outcome of improving student reading ability.
2. Consistency is built into this professional development delivery system. First, the purpose and content of the training is stable because the information pertains to consistent state standards and assessments. Second, stability is maintained because all trainings use the same teacher-made materials. These materials include hard copies and jump drives. Third, classroom videos are posted so teachers can watch their colleagues share content while practicing exemplary pedagogy. Last, emulating the National Writing Project, the train-the-trainer delivery system includes training and shadowing. Novice teacher trainers work together with seasoned trainers, shadowing their experienced peers before assuming the full responsibility of delivering instruction to their peers. In these ways, the delivery of strategies.
3. The training is practical. Though the research behind the strategy is theoretical, the teacher trainers are practical. Educators attend because they want to know the strategy to grow their students. They want to apply what they learn in their classrooms tomorrow and they want to adapt the methods to fit their own styles. The training follows a workshop model which allows teachers time to practice with one another during the weekend. Simultaneously, teacher facilitators model best teaching methods during their delivery.

4. This method is fiscally responsible. A cadre of teacher practitioners faced one-fifth of the state's teachers, driving the long highways to over a dozen towns and cities in all points of the state. The cost of contracting a team of teachers was a fraction of the cost of utilizing other, often out-of-state, vendors who may not know the state standards and assessments, the districts' idiosyncrasies, and the student data.

What content must be delivered?

In *All Systems Go*, Michael Fullan, Ontario's advisor to education and professor emeritus at the University of Toronto, discusses whole system reform. He says that a state educational system must discipline itself to focus its efforts. Ontario schools took his advice, beginning with literacy, and schools saw results (2010). The Wyoming Department of Education, with the mission of improving instructional quality to raise student achievement, took Fullan's lessons to heart. Because the state reading test was termed 'instructionally supportive' by the technical advisory committee composed of national assessment leaders, the mission of the state's professional development was to share the essential learnings in reading that matched state standards and formed the basis of the state test (Popham, 2003).

This reading content is universal, however, and can be used to improve reading performance regardless of the nature of the state test. Expert teachers shared the work of renowned reading expert, Roger Farr, professor emeritus of the University of Indiana (Popham, 2003). Farr's work applies to the new Common Core literacy standards across the contents.

Farr synthesized the world of reading instruction and assessment to eight high-powered cognitive skills. According to Farr, good reading instruction focuses the attention of readers by teaching them what to expect from three different kinds of texts, how those texts can fulfill their purposes for reading, and the strategies for constructing the meaning they need from different texts. To see if students comprehend their texts, teachers ask students to write reader responses, termed constructed responses. These constructed responses are used as a common classroom assessment to review and inform grade level meetings. The essential kinds of texts for students to read and comprehend are: narrative, expository and functional (maps, graphs, brochures, advertisements). Furthermore, each of these texts is read for certain purposes:

1. narrative--to comprehend plot, theme and story elements affecting the story;
2. expository--to comprehend main idea and supporting details, organizational structures, and broad issues;
3. functional--to comprehend relevant information and application of this information.

These three essential texts and the eight purposes for reading them are assessed on the state test through multiple choice and constructed response questions. The weekend warriors focused on the constructed response--how to select texts, how to write prompts for students, how to teach students to write relevant, accurate and sufficient responses, and how to score the student products with inter-rater reliability. After practicing in the weekend workshop, attendees returned to their schools armed with a valid assessment tool to use as a grade level common assessment in their professional learning groups. This giant statewide professional teaching community produced improved scores; but what really matters is that students began to read and comprehend texts better, a life skill that goes far beyond performance on a state test.

Closing

Federal scrutiny on teacher effectiveness is intensifying. Federal and state lawmakers are working bills to hold teachers accountable for student performance on standardized tests. Teachers, then, must have optimal opportunity to learn and grow. The delivery method for professional development that produces evidence of success is peer-to-peer work--at the grade, building and state levels. Teachers prefer to learn with expert colleagues, those who work daily in the classroom and schools, those who have proven success, those whose work is authentic and worthy of emulation, those who have deep content knowledge, and those who artfully model their pedagogy.

According to researcher John Hattie (2009), the quality of the classroom teacher is the most significant indicator of student success. Other characteristics count to make a great school, but ultimately, raising student achievement depends on teachers. Our most precious resource is our children. Teachers need and deserve the best professional development to improve their content knowledge and pedagogy and meet the learning needs of these students. To realize teachers' potential to achieve instructional excellence, professional development must be done with teachers, not to them.

References

Arapahoe school shows student testing improvement. *Billings Gazette*. December 13, 2012.

Fullan, M. (2010). *All systems go: The change imperative for whole system reform*. Thousand Oaks, CA: Corwin.

Brown, J. (March 2, 2012). *Denverpost.com*. "Cost doesn't spell success for Colorado schools using consultants to improve achievement."

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Hattie, J. (2009). *Visible learning*. New York: Routledge.

<http://www.centerforcsri.org/plc/program.html> (information on professional learning communities)

<http://www.nwp.org/cs/public/print/doc/about.csp> (information on National Writing Project)

<http://www.paec.org/teacher2teacher/> (information on T2T initiative by United States Department of Education)

<http://www.readingrecovery.org/reading-recovery/teaching-children/basic-facts>

Popham, W.J. (2003). *Crafting curricular aims for instructionally supportive assessment*. Los Angeles: University of California.

Popham, W. J., Pellegrino, J., Berliner, D. Tests helping teachers. *American School Board Journal*. March 2012.

Scafidi, B. (February 2013). The school staffing surge: Decades of Employment Growth in American's Public Schools. Friedman Foundation. edchoice.org.

PAWS Reading Improvement

| |
|-------------------------------------|
| Domain: PAWS |
| Cube: Grades 3-8 - PAWS Performance |
| |

PAWS - Grades 3-8 - PAWS Performance

| | Reading : Average Scale Score | | | | | |
|-----------------|-------------------------------|---------|---------|---------|---------|---------|
| | 2005-06 | 2006-07 | 2007-08 | 2008-09 | 2010-11 | 2011-12 |
| Grade 03 | 595 | 613 | 592 | 581 | 597 | 599 |
| Grade 04 | 645 | 663 | 657 | 654 | 668 | 675 |
| Grade 05 | 655 | 662 | 647 | 648 | 670 | 670 |
| Grade 06 | 661 | 674 | 670 | 674 | 683 | 687 |
| Grade 07 | 685 | 690 | 675 | 669 | 684 | 688 |
| Grade 08 | 688 | 702 | 693 | 686 | 697 | 700 |

As seen in the above table, reading scores in Wyoming have improved over time as teachers became more focused on reading instruction. PAWS reading is an instructionally supportive instrument. It is a criterion referenced test rather than a standardized one. This means teachers and students teach and learn the criteria and the targets. These criteria are the essential components of reading comprehension, hence, worthy of spending teaching time and reading time.

School-Targeted Improvement: Arapahoe School Project

The Wyoming Department of Education entered into a reading improvement project with Fremont 38 in the summer of 2011. This project created a lab school at the K-8 Arapahoe School with potential expansion and outreach to all WY K-12 schools emphasizing Title I schools in improvement in reading.

The Story

Due to low student performance on the reading portion of the PAWS test, Fremont 38 initiated the project in late July by contacting the WDE for technical assistance. The WDE responded with a visit to the school, meeting with members of the leadership, board and faculty. Attending the meeting from WDE were Superintendent Hill, co-deputies and other members including representation from Wyoming State System of Support and Special Services. At that meeting, Superintendent Hill expressed the need for full support from the Fremont 38 School Board which was granted at a subsequent school board meeting. She also indicated that WDE support would be extended beyond Fremont 38 to all schools in the state who seek improvement and are interested in participating.

What's WYR?

WYR works with struggling students in need of accelerated reading growth. WYR (WY Read) is a Tier III tutorial based on the work of Richard Allington (see www.whatworks.gov) and Richard Allington's meta-analysis in his book *What Really Matters in Rtl*), Marie Clay, and the

National Reading Panel. It has been implemented in Wyoming for approximately 10 years, accelerating student growth from 1.5-4+ years in 50 lessons. WYR is a Tier III intervention.

The application of theory for WYR requires the student to read a self-selected, appropriately difficult book under the watchful eye of a trained tutor who is a careful observer and skilled at explicit teaching and specific reinforcement. The protocol for the tutorial is based on the reading elements explicated by the National Reading Panel. These sub skills are taught within the meaning-rich text the student is motivated to read, and the tutor seeks to bond with the student while offering highly effective, scaffolded support. This intervention is daily for 30 minutes for approximately 50 lessons.

Working shoulder-to-shoulder with students, parents, staff, and WDE, the focus on literacy produced positive results. For example, the cohort of students performing at 18% proficient/advanced in third grade in 2011 improved to over 58% in fourth grade in 2012. All other grades improved significantly as well. This kind of growth occurs when, as Elmore’s says: “We learn to do the work by doing the work, not by telling other people to do the work...”

Arapaho Reading Scores Show Growth in One Year

| Grade | 2011 Percent Proficient and Advanced | 2012 Percent Proficient and Advanced | % Change |
|-------|--------------------------------------|--------------------------------------|----------|
| 03 | 18.4 | 34.2 | 15.7 |
| 04 | 47.2 | 58.5 | 11.3 |
| 05 | 33.3 | 60.5 | 27.2 |
| 06 | 64.0 | 64.3 | 0.3 |
| 07 | 46.7 | 59.3 | 12.6 |
| 08 | 47.4 | 62.1 | 14.7 |

As noted in the above table, Arapaho student reading scores show significant improvement from 2011 to 2012 due in part to the one-on-one intervention focused on comprehension. This school also fully implemented 3+8 reading.

References

Allington, R.L. (2012). *What really matters for struggling readers: Designing research-based programs.* (3rd ed.). Boston: Allyn & Bacon.

Clay, M. (1993). *Reading Recovery.* (1st ed.). Portsmouth: Heinemann.

Hattie, J. (2012). *Visible learning.* New York: Routledge.

National Reading Panel. (1998). *Preventing reading difficulties in young children*. Washington D.C.: National Academy Press.

What Works Clearinghouse. (2007). Find what works: Summarize and compare the evidence of the effectiveness of interventions that address your school or district's needs.

Retrieved January 30, 2013, from

www.iew.ed.gov/ncee/wwc/findwhatworks.aspx#Gneral%20literacy%20achievement.

May 13, 2013

TODAY'S TEACHER EVALUATION: A CAUTIONARY NOTE

by

W. James Popham, and Rick Stiggins*

We write because we worry. Our worries stem from today's expansion of high-stakes teacher-evaluation programs all across America. We fear, because of the meager knowledge-base regarding how best to evaluate teachers, some ill-conceived or poorly implemented teacher-evaluation programs will lead to an unnecessarily large number of inaccurately evaluated teachers.

Unsound teacher-evaluation programs, of course, harm mistakenly evaluated teachers. This is particularly apparent when such high-stakes decisions as tenure-denial and job-termination are on the line. More importantly, flawed teacher-evaluation programs can also harm students. If genuinely *effective* teachers are mistakenly judged to be ineffective, those teachers—in an effort to “improve”—may abandon instructional procedures that are working just fine. Conversely, *ineffective* teachers mistakenly identified as satisfactory won't receive assistance they need to improve. In both instances, because of inaccurate teacher evaluations, many students end up receiving less effective instruction.

The recent surge of teacher-evaluation programs, of course, can be traced to two federal school-reform initiatives, both of which call for states to create substantially tougher teacher-evaluation procedures. The 2009 Race to the Top Program promised substantial federal dollars to states where, among other reforms, stronger teacher-evaluation procedures were installed. Two years later, a federal waiver program permitted states to dodge serious sanctions associated with the No Child Left Behind Act by strengthening their teacher evaluation systems. These two federal incentives, then, spurred all but a few states to overhaul or replace their existing teacher-evaluation programs. Such teacher-evaluation procedures were fashioned either by a state's legislators or education officials in an attempt to position their state to receive federal approval.

The new federal guidelines call for teacher evaluators to employ multiple kinds of evidence—one of which is to be student growth usually represented by students' test scores. Moreover, growth-evidence must function as a “significant” evaluative factor. Another key guideline calls for a state's teacher-evaluation system to “inform personnel decisions.” Clearly, federal authorities were advocating a robust process in which ineffective teachers could lose their jobs

or be denied tenure. With such high stakes, of course, only the highest quality evidence should be employed.

Because there is, as yet, scant research evidence *or* time-tested practical experience regarding how to evaluate teachers accurately, we urge caution when implementing recently devised teacher-evaluation procedures, even those designed by well-intentioned state personnel or fashioned by well-intentioned state legislators. Good intentions aside, some of the resultant evaluative procedures embody serious shortcomings.

For example, most teacher-evaluation programs rely heavily on observations of classroom-practice. If those observations employ a research-based and experience-honed observation system, use well-trained and certified observers, and collect evidence from a sufficient number of class sessions, then such evidence can make an important contribution to the evaluation of a teacher's effectiveness. On the other hand, the use of slap-dash observational schemes employed by untrained observers who only look at a handful of atypical "showcase" lessons will provide misleading evidence of a teacher's skill.

Another instance of an evaluative mistake often made these days involves a state's use of inappropriate tests to assess student growth. Many states currently measure student growth by using students' scores on annual accountability tests. Such states sometimes employ incomprehensible statistical gyrations to provide a "value-added" estimate of how much a teacher's students have learned—an estimate that usually fails to rule out other contributors to students' test scores. Yet, most tests currently being used to collect this evidence are not accompanied by evidence indicating that those tests are able to distinguish between well taught and poorly taught students. Any test without at least some evidence that the test can differentiate between strong and weak instruction should *never* be used to evaluate teachers.

Teachers' classroom tests, if properly constructed, appropriately administered, and scored accurately, can promote students' learning and also provide powerful evidence of students' growth toward officially approved curricular goals. Thus, such assessments can contribute to the teacher-evaluation process, but only *if* teachers and their supervisors are given the opportunity to learn how classroom assessments can provide evaluative evidence that is both accurate *and* credible.

Finally, the use of students' ratings to help evaluate teachers is beginning to be incorporated in numerous teacher-evaluation programs. Yet, while one recent study indicates that student perceptions of a teacher's skill correlate with other measures of learning, it is not yet known whether students' responses will be distorted when students know their ratings will be routinely used to evaluate teachers. Thus, until the kinks have been ironed out of this potentially rich source of evidence, its evaluative weight should be modest.

Appendix C

Summing up, both of us have been working with teachers and school leaders for many decades. We recognize the potential benefits, for both students and teachers, of a rigorous teacher-evaluation process—the kind contemplated in recent federal initiatives. But many educational leaders think more is known about how to evaluate teachers than is, in fact, the case. Accordingly, we urge considerable caution when applying any recently birthed teacher-evaluation system. Such procedures should definitely be piloted before widespread implementation—and when their anticipatable shortcomings are identified, revisions should be implemented immediately. In short, we recommend a consummately cautious approach when evaluating a state’s teachers.

*W. James Popham is a UCLA emeritus professor and Rick Stiggins is the founder of the Assessment Training Institute in Portland, Oregon. Both have written extensively on instruction, assessment, and evaluation.

The Federal Program *Reading First* Caused Unintended Consequences
By Sheryl Lain

With the federal enactment of No Child Left Behind in 2001, state and federal lawmakers began to emphasize early literacy instruction. The WDE strategic plan made K-3 reading a goal, claiming that children who do not learn to read by third grade are destined to be reading deficient all through school.

However, the state government's focus on K-3 reading literacy did not produce the desired result. The PAWS results for third grade reading lag behind math scores. The third grade reading scores are lower than in subsequent elementary grades.

The federal government's focus on early literacy did not produce the desired results either. The federal government spent a billion dollars a year over a five-year period on the early literacy program known as Reading First. According to NAEP, only 31% of the nation's fourth graders performed at or above proficient in reading in 2005. The results were only slightly higher in 2007.

Wyoming lawmakers complied with the federal focus on early literacy by requiring a new data collection (Data Collection Form 626) wherein each district was required to list the textbook, tests, and programs they use in grades K-3 including the programs and tests they use in interventions. The emphasis was on purchasing materials not on growing teachers. One unintended consequence was districts purchased a proliferation of programs and tests (see 626 report for vast array of products and third grade scores). Another possible consequence is lower than necessary PAWS reading scores in third grade due to the mismatch of these programs and tests with the PAWS test which emphasizes comprehension.

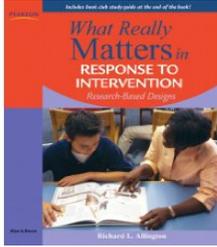
Simultaneously, WDE encouraged the use of programs and tests. For example, the special education division hired a consultant who encouraged the use of an SRA product known as Reading Mastery, which was then purchased by thirty districts. The WDE also sponsored trainings on the use of textbooks and assessment such as DIBELS, which measures how fast students can read. To comply with a Bridges Summer School requirement, districts bought NWEA MAP, a computer adaptive assessment, to see if summer school students' reading improved. An unintended consequence of the required purchase of MAP is the proliferation of this assessment which is not advocated by assessment experts.

These commercial textbooks, programs and assessments are not aligned to state standards and PAWS. Also, these products are often "scripted" meaning that the teacher must read or recite the teacher's guide word for word to children during the lessons. Trainers insisted that teachers follow the script, so instruction is "teacher proof." Often districts purchased materials, hired consultants to train teachers to use the materials and used Instructional Facilitators to monitor the teachers' use of the materials.

In spite of the state legislature's focus on early literacy and the districts' compliance, the results do not demonstrate that the time and money was well spent. Third grade scores fluctuated widely over the past two years alone. Forty-two districts improved in 2011-12 and one year later, twenty-nine districts declined. One inference can be made: using core programs does not guarantee student success.

What does research say about “teacher-proof” scripted programs? This question among others should be answered in further studies. Two well-known researchers studying a vast array of studies decry these scripted approaches calling instead for growing teacher experts. The work of Linda Darling-Hammond and John Hattie is explained elsewhere in the report.

A synopsis of a meta-analysis on reading, from Richard Allington’s book, *What Really Matters for Struggling Readers*, is on the next page. Allington says that teaching reading requires expert teachers who know how to adapt teaching for each individual learner. In his book, Allington mentions What Works Clearinghouse, a list of programs commonly purchased in the United States. See the list from the Clearinghouse at the end of the Allington synopsis. The reader will notice only one “program” that is successful. This “Program” is Reading Recovery which is a staff development model not a commercial textbook program.



What Really Matters in Response to Intervention
By Richard Allington
A meta-analysis of 900 studies+

Chapter 1

- Materials must be appropriate for learners, at the student's instructional level.
- **Special education** students don't catch up with their peers. We need to **double and triple reading growth** by adding the gift of **time** with kids taught by a **highly skilled teacher**.
- Doing nothing is a good strategy for fostering **dropping-out-of-school behavior**.
- We need classrooms where readers **get more time and more intensive reading lessons** targeted to their specific needs.
- **How much time?** Besides normal classroom time of 1.5 hours per day at least, a daily 30-minute tutorial.
- "We don't need to hire a consultant to determine why subgroups fail to make AYP. If a school has classrooms with books on kids' desks they cannot read, what can we expect but low or no growth?"

Chapter 2

- Too often special education classes use textbooks kids can't read.
- The more whole-class teaching, the lower the academic achievement.
- A supplementary reading program must be designed consistent with the evidence. Texts must be appropriately difficult.
- Cycling students through packaged programs that are not responsive to the individual needs...has the potential to lead to more kids identified as learning disabled rather than fewer."

Chapter 3

- Recommended oral reading accuracy levels of text difficulty:
Independent reading = 99-100%

Instructional reading = 95-98%

Frustration reading = avoid

- Kids need to read high success books to build reading stamina.
- “Whenever we design an intervention for strugglers that produces success, we match them with texts they can read with a high level of accuracy, fluency and comprehension.”

Chapter 4

- To produce growth, strugglers need to be engaged in reading high-success texts for roughly **two-thirds** of the intervention time.
- Dramatic **increase in reading volume** produces results.
- During reading they **practice all those complicated strategies and skills they are developing in unison**.
- **Rereading** is key also.
- In a **30 minute tutorial, kids need to spend 20 minutes in high-success reading: 5 minutes each for word work and thinking strategies**.
- We need access to lots of books that kids can read accurately, fluently and with understanding.

Chapter 5

- **A meta-analysis** using over 900 studies shows that 1:1 or very small groups (1:2) is key to accelerate reading growth. However, in Title I and Special Ed, group size is more like 1:5-9, which accounts for why strugglers only gain 2-3 months of growth per year.
- Do not pull strugglers for reading tutorials from classes where they read. These kids need extended time.
- Use an innovative local design. “I argue that, without spending any more money, virtually every school and district has sufficient adults to provide all of the needed tutorial and very small group reading interventions. There are two adults for each classroom teacher and most employees do not deliver instruction.”

Chapter 6

- Curricular coordination is when two curricula support the same philosophies of reading with similar strategy use. What is learned in one setting is transferable to another.
- Most strugglers receive interventions not informed by classroom reading lessons. More interventions are in curricular conflict than not.

Chapter 7

- It is more common to see strugglers complete workbook skills with an inexpert para than to see strugglers reading with an expert.
- Buildings should spend “less time buying commercial products. As noted by What Works Clearinghouse, there are no core reading programs that have adequate evidence” of effectiveness.” Only Reading Recovery has high effectiveness and it is a teacher training program. www.whatworks.ed.gov

Chapter 8

- Expert teachers seem to get results because they are better at “matching struggling readers with appropriate texts;” they adapt more quickly; their students do more actual reading and writing.
- “Interventions focused on developing meta-cognition and comprehension repeatedly show more reading growth than decoding emphasis.” And this pertains to low and high poverty schools.
- This is because teachers: maximize time to read; focus on meaning; provide discussion; integrate reading and writing with content.
- The research available indicates that meaning-emphasis classroom reading instruction produces significantly better reading achievement than skills-emphasis instruction.

Chapter 9

- Classroom libraries need: different levels, displays, student choice.
- Classrooms need: authentic reading/writing tasks, teacher reinforcement; higher order thinking; connections; expressive read alouds; small group instruction; teacher modeling; minilessons; smooth transitions; positive classroom management.”

- The Rosalie Fink study says: help kids find their passionate interest that sustains independent reading; provide them lots of reading on this passionate topic; support their schema knowledge about the topic; teach them powerful strategies as they read interesting texts; provide mentoring support.

Chapter 10

- Commercial products don't work as interventions
- No commercial program received high ratings across all four proficiency reading areas
- Reading Recovery is the closest and it's more a training program
- The best assessments are running records and words correct per minute. Also Informal Reading Inventories are good for initial screening and annual reviews
- Most computer intervention programs are not very effective at raising reading achievement because they are worksheets on computer

What Works Clearinghouse - Review of Reading Recovery

This chart shows that Reading Recovery outperforms all other programs in the key reading areas of alphabetic, fluency, comprehension and general reading.

| WHAT WORKS CLEARINGHOUSE | | | | |
|--|--------------|-----------------|---------------|-----------------------------|
| IMPROVEMENT INDEX | | | | |
| INTERVENTION NAME | ALPHABETICS | READING FLUENCY | COMPREHENSION | GENERAL READING ACHIEVEMENT |
| Accelerated Reader | NA | +3* | 0** | +16* |
| Cooperative Integrated Reading and Composition (CIRC) [®] | NA | NA | +4** | NA |
| Corrective Reading | +9* | +11* | +7* | NA |
| Classwide Peer Tutoring [®] | NA | NA | NA | +14* |
| DaisyQuest | +23* | NA | NA | NA |
| Early Intervention in Reading (EIR) [®] | +36* | NA | +18* | NA |
| Earobics [®] | +25* | +15* | NA | NA |
| Failure Free Reading | +1* | +2* | +10* | NA |
| Fast ForWord [®] | +8* | NA | +1* | NA |
| Fluency Formula™ | NA | +10* | -11* | NA |
| Ladders to Literacy | +25** | +26* | +9** | NA |
| Lexia Reading | +11* | +9* | +11* | +9* |
| Lindamood Phonemic Sequencing (LiPS) [®] | +17* | NA | +6* | NA |
| Little Books | NA | NA | NA | +12* |
| Peer-Assisted Learning Strategies (PALS) [®] | +19* | +13* | +13* | NA |
| Read Naturally [®] | NA | +8* | +2* | NA |
| Read, Write, and Type! [™] | +8* | NA | +3* | NA |
| Reading Recovery[®] | +34** | +46* | +14* | +32** |
| Sound Partners | +21** | +19** | +21** | +9* |
| Spell Read | +18* | +9* | +20* | NA |
| Start Making a Reader Today [®] (SMART [®]) | +16* | +17* | +14* | NA |
| Stepping Stones to Literacy | +30* | NA | NA | NA |
| Success for All [®] | +13** | NA | +8** | +10** |
| Voyager Universal Literacy System [®] | +11** | NA | -25* | NA |
| Waterford Early Reading Program | +19* | NA | +4* | NA |
| Wilson Reading | +13* | +6* | +7* | NA |

NOTE: Average percentile points refer to the difference between the percentile rank of the average treatment student compared to the percentile rank of the average control student. Extent of evidence categorization: *small; **medium to large. Evidence is categorized as small if, for a given outcome domain, it is based on only one study, or from only one school, or from a total sample size of less than 350 and a total of less than 14 classrooms across studies. Evidence is considered medium to large extent for a given domain outcome if it is based on more than one study, and from more than one school, and the total sample size is at least 350 students or from at least 14 classrooms across studies. NA = not applicable (Source: WWC, October 2011)

Learn how Reading Recovery can provide lasting results at ReadingRecoveryWorks.org