

Rereading Fluency

Process, Practice, and Policy

Bess Altwerger
Nancy Jordan
Nancy Rankie Shelton

Foreword by Richard L. Allington

HEINEMANN
Portsmouth, NH

Heinemann

A division of Reed Elsevier Inc.
361 Hanover Street
Portsmouth, NH 03801-3912
www.heinemann.com

Offices and agents throughout the world

© 2007 by Bess Altwerger, Nancy Jordan, and Nancy Rankie Shelton

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without permission in writing from the publisher, except by a reviewer, who may quote brief passages in a review.

Library of Congress Cataloging-in-Publication Data

Altwerger, Bess.

Rereading fluency : process, practice, and policy / Bess Altwerger, Nancy Jordan, Nancy Rankie Shelton ; foreword by Richard L. Allington.

p. cm.

Includes bibliographical references and index.

ISBN-13: 978-0-325-01034-2

ISBN-10: 0-325-01034-X

1. Reading comprehension. 2. Fluency (Language learning). I. Jordan, Nancy C. II. Shelton, Nancy Rankie. III. Title.

LB1050.45.A45 2007

428.4'3—dc22

2007025644

Editor: Leigh Peake

Production management: Appingo

Production coordination: Vicki Kasabian

Cover design: Night & Day Design

Typesetter: Appingo

Manufacturing: Louise Richardson

Printed in the United States of America on acid-free paper

11 10 09 08 07 VP 1 2 3 4 5

Dedicated to the teachers and students who
struggle each day to find meaning, purpose, and joy in literacy
learning despite the many challenges that have been
placed in their way. They are our heroes.

Contents

Foreword by Richard L. Allington	vii
Acknowledgments	xi
PART I: WHY ALL THE FUSS OVER FLUENCY?	
1 Why Do We Need a Critical Look at Fluency?	1
2 What Is Fluency and How Important Is It?	9
3 Following the Fluency Trail: From the National Reading Panel to Reading First	19
PART II: WHAT DO WE KNOW ABOUT FLUENCY AND THE READING PROCESS?	
4 How Can Young Readers Inform Us About Fluency?	27
5 What Is the Relationship Between Fluency and Comprehension?	37
6 What Is the Relationship Between Fluency and Decoding?	44
7 Phonics Versus Literature Programs: What's the Difference?	53
8 Does Fluency Distinguish Between More and Less Proficient Readers?	63
9 What Does DIBELS Tell Us About Readers?	73

Part III: RETHINKING FLUENCY

10	Should Fluency Be Considered a Critical Component of Reading?	88
11	If Not Fluency, Then What?	97
12	Should Fluency Be Used to Make Instructional Decisions?	103
	References	110
	Index	116

Foreword

Rereading Fluency: Process, Practice, and Policy is an important and timely book. It is important because the authors present a number of challenges to the dominant model of reading instruction and development—challenges raised from the analyses of the data they gathered in schools and classrooms.

Their challenges are far from the only ones raised about the dominant paradigm. The data-based findings presented here are reinforced by a variety of studies already available. But most of these studies attended to only a single challenge to the current implementation of “evidence-based” reading instruction. Altwerger, Jordan, and Shelton take a broader view.

Nonetheless, let me note just a few of the studies that provide good evidence, converging evidence, that the challenges presented in this book need serious consideration from teachers, administrators, and policy makers.

Challenge: Good readers are fast readers. Walczyk and Griffith-Ross (2007), Pressley (2006), and Buly and Valencia (2002) argue that this simply is not true. Walczyk and Griffith-Ross report that “slowing down” is the most commonly used strategy when readers encounter difficulty. The second strategy is rereading. Both result in slower reading rates. Both result in better comprehension. Pressley found that “children can often read with great speed and accuracy and yet recall few of the ideas in the text they read” (2006, 209). In a similar vein, Buly and Valencia found that nearly 20 percent of their struggling readers could read fast and accurately but exhibit almost no recall or understanding of what they read. While focusing instruction on reading faster may generate observed gains on rate/fluency measures, such gains do not indicate that children are improving as readers.

Challenge: Reading rate and accuracy measures can be used to develop better instruction. Schilling and others (2007) found that 37

percent of the third-grade students who were identified as at low risk on the DIBELS oral reading fluency assessment fell below grade level reading achievement on the Iowa Test of Basic Skills administered at the end of the school year. Because their DIBELS benchmark scores indicated little or no risk of reading failure, many of these students were denied access to instruction that might have resolved the risk that wasn't evident on the DIBELS tests.

Challenge: Commercial core reading programs will improve reading achievement. In a large-scale analysis across thirty-seven school districts, McGill-Franzen and others (2006) found that in schools that used one of the state-approved commercial core reading programs to guide instruction, one-third of the third-grade students, on average, failed to achieve the minimum benchmark on the state reading assessment. In all these schools the DIBELS assessment was also being used to inform instruction. But still large numbers of children failed to learn to read. So much for the science of fluency assessment and core reading program requirements.

I could go on, but I hope that I have made my point: The findings reported in *Rereading Fluency: Process, Practice, and Policy* are supported by other scientific research. It is the scope of the work reported in this book that sets it apart. Additionally, the authors do not just criticize current policies and practices but offer alternatives for improving the quality of reading assessment and instruction.

This book challenges the conventional view of how to improve reading instruction and achievement. It is through such challenges that science ultimately finds its way to fact.

Richard L. Allington, Ph.D.
University of Tennessee

References

- Buly, M. R., and S. W. Valencia. 2002. "Below the Bar: Profiles of Students Who Fail State Reading Assessments." *Educational Evaluation and Policy Analysis* 24 (3): 219–39.
- McGill-Franzen, A., C. Zmach, K. Solic, and J. L. Zeig. 2006. "The Confluence of Two Policy Mandates: Core Reading Programs and Third-Grade Retention in Florida." *Elementary School Journal* 107 (1): 67–91.
- Pressley, M. P. 2006. *Reading Instruction That Works*, 3d ed. New York: Guilford.
- Schilling, S. G., J. F. Carlisle, S. E. Scott, and J. Zeng. 2007. "Are Fluency Measures Accurate Predictors of Reading Achievement?" *Elementary School Journal* 107 (5): 429–48.
- Walczyk, J. A., and D. A. Griffith-Ross. 2007. "How Important Is Reading Skill Fluency for Comprehension?" *Reading Teacher* 60 (6): 560–69.

4 How Can Young Readers Inform Us About Fluency?

In our roles as classroom teachers, reading specialists, and university professors over the years, we have had numerous opportunities to observe children engaged in school reading activities. Since the assessment of student learning has been central to our teaching and curricular decisions, we have devoted much of our professional lives to studying and using a wide range of literacy assessments including miscue analysis, running record, and curriculum-based assessments. Our considerable experience and knowledge have taught us the limited value of standardized test scores in helping us understand our students as readers and writers.

In recent years, we have witnessed the rise to prominence of fluency instruction and assessment in schools under the Reading First guidelines of No Child Left Behind (NCLB). In many schools, one-minute fluency assessments have replaced other analyses of student reading (close observation, miscue analysis, running record, work sampling, and parent feedback) that we consider more valuable for understanding and responding to students' unique reading needs. Yet we have found a lack of independent research providing scientific evidence (beyond standardized test scores) for this shift toward fluency assessment. As supervisors of student interns in area schools, we have also observed classrooms where, due to the fluency instruction and assessment craze, children are pressured to read faster and faster. Along with our student interns, we became concerned about the potential consequences of fluency-boosting practices our interns were expected to implement in their school placements. We realized that it was critical, not just for theoretical reasons but for current and future students and teachers, to more firmly examine the role of fluency in the reading process and in reading development.

But how could this be accomplished in a climate where mandates attached to funding have silenced or ignored many voices, including our own, and where debate and inquiry have been curtailed? As teachers and scholars, we felt that our best option was to conduct research. We strongly believed that claims made by politicians, consultants, textbook companies, administrators, and the director of National Institute of Child Health and Human Development (NICHD) about the importance of fluency needed to be subjected to scientific examination. Therefore we set up our study to conduct that examination and provide the field with further data that would help to assess the validity of these claims. We hoped to use our findings to inform our own teaching and to inform others who are impacted by current federal policies on fluency. And so we embarked on our journey.

Fortunately, we had already generated a wealth of data that provided us with the opportunity to conduct our inquiry. As members of a team of literacy researchers at Towson University, we had developed a study to respond to the proliferation of scripted, skill-based reading programs that were being mandated across our state under new system policies. We designed a research project to examine the impact of four contrasting reading programs, two of which represented the commercial programs that were being mandated and two of which were non-commercial literature-based programs. More specifically, these four reading programs were McGraw-Hill's SRA Reading Mastery [Direct Instruction] (Engelmann et al. 1995)]; McGraw-Hill's Open Court (2000); a county-developed literature-based reading program; and a guided reading program loosely based upon Fountas and Pinnell's model (1996). The primary levels of both McGraw-Hill programs focused predominantly on systematic phonics and skills for the teaching of reading. Both the literature-based and guided reading programs approach reading instruction as a holistic, meaning-based process. We hoped that our research would provide us with urgently needed information regarding the impact of contrasting reading programs on students' development of the reading process, including their comprehension, strategies, and the use of phonics both in and out of context. (See Altwerger et al. 2004, Altwerger 2005, and Wiltz and Wilson 2005 for further information on our findings.)

The theoretical tenet guiding our research design was that oral and written language are complex and socially constructed systems that

cannot be fully understood without consideration of their intricate relationships to situated contexts. This intricacy and complexity is not taken into account in reading research based solely in quantitative experimental design. Therefore we created a methodology that would enable us to achieve an in-depth and multifaceted investigation of how students in various instructional programs develop as readers. Accordingly, we examined the impact of these four different reading programs using both quantitative and qualitative methods for data collection and analysis.

Approximately thirty children from each of the four program sites participated in the study. We selected a stratified random sample from all potential participants that teachers and/or schools identified as high, middle, or low readers. This enabled us to investigate relationships and distinctions among groups of students considered to be at varied levels of reading competence within and across program sites. This also helped us to examine the validity of grouping and labeling readers using running record accuracy scores and Open Court and Direct Instruction assessments. All of the participants were carefully screened to ensure that they had been taught using their respective reading program for a minimum of two years, were proficient English speakers, and received no special education services. Also, the children participating in the study were matched across school sites for percentages of free and reduced lunch (87 to 100 percent).

For the qualitative component of our study, we included classroom observations, literacy artifacts, and interviews with students and teachers. Our intent was to describe in detail the ecology of the classrooms so that we would be able to draw comparisons across programs in ways that gave us some understanding of how students construct knowledge about texts and the process of reading.

Our quantitative data included reading samples collected from each participant. Given that we hoped to generate representative samples of children's oral reading and comprehension of challenging, but not frustrating texts, we carefully selected books that spanned eight reading levels (kindergarten through fourth grade) based upon the gradient developed by Fountas and Pinnell's (1996, 2001) lists of leveled books. Each book had an identifiable story structure. We included three books at each level to be sure we had at least one selection that would be unfamiliar to each reader. Students were offered the selection of books

that were deemed to be on or close to their instructional levels (based upon teacher recommendation and researcher judgment formed in the initial stages of the research sessions). Following miscue analysis protocol (Goodman, Watson, and Burke 1987), each student was then asked to read the target book aloud (without researcher assistance) and provide an unaided and aided oral retelling of the story. Each reading and retelling event was audiotaped to ensure accuracy and allow for reliable and precise coding of data. For each of these oral readings, we conducted a complete miscue analysis.

As discussed in Chapter 2, a miscue is defined as a reader's deviation from a written text while reading aloud. Miscue analysis generates an assessment of these deviations in order to document how readers integrate the cueing systems: graphophonic, the set of relationships between oral and written language; syntactic, the grammatical interrelationships of words and sentences; and semantic, the meanings of words and text. For our study, the researchers listened closely to students' oral readings during the research sessions, noting miscues and other pertinent behaviors on their printed copies of the stories (worksheets). We later returned to the audiotaped version of the readings to more completely mark the worksheets. For example, in Figure 4-1 the reader uses repetition (repeats what has already been read) as she reads. These parts are underlined and marked with (R) and are not considered miscues. In line 07 04 the reader read *they* for *I* and then self-corrected. The worksheet was marked with *I* underlined and *they* written above it and then marked as a self-correction (C) (see Figure 4-1).

From these marks, a coding sheet was completed which analyzed the miscue in order to note the reader's use of cueing systems and strategies. Another researcher independently listened to the tape and coded the miscues. Interrater reliability was established for the coding of miscues at .90. Miscue data was analyzed for syntactic acceptability, semantic acceptability, meaning change, correction, meaning construction, grammatical relationship, sound similarity, and graphic similarity. For each miscue, researchers rated the extent to which the miscues reflected these variables. For example, a substitution miscue could have HIGH, SOME, or NO graphic or sound similarity to the text word; that miscue could also be fully semantically or syntactically acceptable (Y), partially acceptable (P), or not acceptable at all (N). From the codings of each miscue for each reader, percentages were calculated (see Figure 4-2).

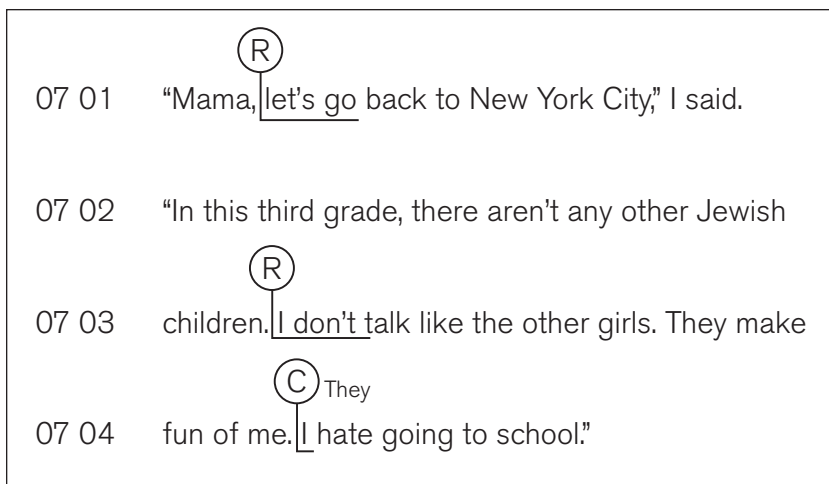


FIGURE 4-1. *Example of Miscue*

For the retelling analysis, a retelling scoring sheet was created for each story. Students' retellings were scored independently by two researchers, through first an unaided retelling and then an aided retelling. Retelling scores were determined based on the identification of setting, characters, theme, and plot episodes; on retelling cohesion; and on the inferences/connections that were made. Again, raw scores were used to calculate a percentage for each retelling.

Advocates of phonics-based programs claim that these programs are better at teaching phonics skills to decode words in isolation and to decode words in extended text. We decided to test this claim. Therefore, in addition to the miscue analysis that assesses the use of graphophonic cues while reading authentic texts, we also administered the Woodcock Johnson Psycho-Educational Battery-R (WJPE-R) Word Attack subtest (Woodcock and Bonner Johnson 1990). This provided a standardized assessment of the children's ability to apply phonics knowledge in isolation.

The data from the miscue analysis and the phonics test were then analyzed using multivariate statistics. Findings were examined in relation to the qualitative data collected at each program site (classroom observations, literacy artifacts, and student and teacher interviews), thereby providing us with a multidimensional and comprehensive picture of the children as readers.

Our observations, literacy artifacts, and interviews confirmed that the teachers had implemented the respective programs. The Direct Instruction and Open Court programs resembled each other in that both devoted time each day to systematic, explicit phonics instruction, used books with stories or texts that progress from one level to the next, and provided lessons that established an instructional ritual: “Children sit through hours and hours of these repetitive reading lessons—new words, different stories, the same chant” (Jordan 2005). The Direct Instruction program, written by the publisher’s in-house authors, controls the vocabulary that is introduced. Each new lesson builds on the one before it by introducing new words and utilizing previously taught phonic elements. Children answer questions in unison or individually. The questions can most often be answered with one word. For example, “What is the name of the store?” And there is always a snap and a recitation. Speed and accuracy are drilled. The Open Court program differs from the Direct Instruction program in that some time is also allocated to teach children to ask questions, to predict and confirm, to use background knowledge, and to make inferences as they read. Nevertheless, there is still only one correct answer, and often that requires just a single word response.

Like the two commercial phonics-based programs, the programs using literature for instruction (guided reading program and literature-based program) resembled each other in important ways. First, the reading lessons focused on constructing meaning from text. Children were taught and encouraged to use multiple language cues, including graphophonics, in order to comprehend while reading. Read-alouds, guided reading lessons (Fountas and Pinnell), the reading of authentic literature, and literature discussions were part of the daily plan. In addition, both programs had a writing component. These programs differed in that the guided reading program relied heavily upon leveled books (Fountas and Pinnell) for most of the instruction, whereas the literature-based program made children’s self-selection of reading material central to the program. At times, the teacher did select multiple copies of authentic texts for guided reading lessons. Other times, groups’ books were selected by the children themselves. While the children in the guided reading program were given topics and structures for their writing, the teachers in the literature-based program used a process approach for the teaching of writing which permitted students to choose their own topics and forms for their writing.

The most significant difference between these two literature-based programs was the amount of time children actually spent reading and writing. In this particular version of a guided reading program, children read connected text an average of only four minutes and fifteen seconds each day (over the course of our observation period). This included time spent reading during their guided reading lessons with the teacher and independent reading tasks. They wrote connected text for even less time. The rest of the language arts block consisted mostly of teacher-directed lessons dealing with how to use language cues to read. In contrast, teachers in the literature-based program devoted at least thirty minutes a day to children's silent independent reading and gave at least another thirty minutes to children for writing. In this program teachers took cues from their students and would allow more time for silent reading and for writing when children seemed actively engaged.

Research Extended to Investigate Fluency

With the above study complete (see Chapter 7 for a discussion of the results), members of the original research team divided into subgroups to focus on different areas of inquiry raised by our first set of findings. As criticisms of the National Reading Panel report's conclusions began to build, including those related to phonics, fluency, and comprehension, (Allington 2002; Coles 2003; Garan 2001; Graves 2001; Krashen 2001; Yatvin 2002), and as our concerns over changes to observed reading practices increased, we decided to make fluency the focus of our next inquiry. We turned to our initial data set, which we recognized as containing the resources we needed (audiotaped readings and retellings) to begin our investigation into the role fluency plays in the reading process.

After a thorough review of the research on fluency, both historical and current (see Chapter 2), we planned a study that would investigate fluency as it is widely being operationalized in policy and practice: as *oral reading rate and accuracy*. We posed the following questions:

1. What is the relationship between reading rate and accuracy and use of graphophonics in and out of context?
2. What is the relationship between reading rate and accuracy and meaning construction during reading?

3. What is the relationship between reading rate and accuracy and comprehension of text?
4. What is the relationship of reading rate and accuracy to meaning and comprehension for readers of high and low proficiency?
5. Do currently mandated commercial phonics-based programs develop reading rate and accuracy better than literature-based programs?

To address these questions, we analyzed all of the audiotaped oral readings accessible to us from the original study (108 oral readings) using four fluency assessment variables we developed specifically for this study. These included words per minute (WPM), or the number of words read during the first minute of reading; words correct per minute (WCPM), or the number of words read correctly during the first minute of reading (currently the most common measure of fluency); WPM_{mis}, or the words read per minute for the text portion in which miscues had been analyzed (enabling us to examine rate and accuracy in relation to other miscue variables); and WPM_{text}, or the words read per minute for the entire story (allowing us to note any discrepancies between fluency results for the miscue section and the first minute).

Interrater reliability was established by having two researchers independently time each reading, with any differences resolved by a third and joint timing. For the data analysis, Pearson Correlation coefficients were generated to examine the relationships among measures of reading rate and accuracy, miscue analysis variables, and the phonics scores. MANOVA was used to examine differences among various groups of readers, such as high and low retellers, on measures of rate, accuracy, strategy use, and comprehension.

Fluency Assessment: DIBELS

Currently, one of the most widely used forms of early literacy assessment is the Dynamic Indicators of Basic Early Literacy Skills, or DIBELS (Good and Kaminski 2002). DIBELS claims to be a complete literacy assessment tool that can be used to make instructional recommendations. In our work with teachers in schools across the nation, we have found that students are being “DIBELED” at an escalating rate. In fact, it seems that the DIBELS Oral Reading Fluency (DORF) test has become one of the most influential programs for grouping and planning

instruction for students. Therefore, we extended our research to include a case-study model of one additional classroom that reflected current practice related to fluency and which used the DIBELS assessment. For this part of our research we asked the following questions:

1. What is the relationship between DORF's measures of rate and accuracy to the same student's rate and accuracy of reading trade books?
2. What is the relationship between DORF and comprehension of literature texts?

As we explained earlier, several small inquiry groups formed from the larger group of researchers in order to investigate specific aspects of the reading process. One of these groups investigating text features (Martens, Arya, Wilson, and Jin 2005) collected new data that replicated our earlier study, audiotaping a group of readers at the end of their second-grade year and conducting a miscue analysis complete with retelling protocol. Since their research focus was on the impact of text features on comprehension, they increased the number of texts read by each participant in the study. Thus, their data set included miscue analyses of multiple readings and retellings for twenty-three readers.

Using this set of audiotaped readers, we replicated our fluency data collection techniques described above, generating rate and accuracy scores for thirteen additional readers for whom DIBELS Oral Reading Fluency (DORF) and Retelling Fluency (RTF) scores were made available. Our findings are reported in Chapter 9.

As we have described, our fluency study is uniquely positioned to provide us with needed information regarding the relationships among various aspects of the reading process, including reading rate, accuracy, comprehension, and the use of phonics both in and out of context. What then did second graders teach us about fluency when it is defined as rate and accuracy? Our findings provide powerful answers to the critical questions we should all be asking about fluency.



DEDICATED TO TEACHERS

Thank you for sampling this
resource.

For more information or to
purchase, please visit
Heinemann by clicking the link
below:

<http://www.heinemann.com/products/E01034.aspx>

Use of this material is solely for
individual, noncommercial use and is
for informational purposes only.