

Closing the Achievement Gap in Ohio with Reading Recovery

Emily M. Rodgers, The Ohio State University
Francisco Gómez-Bellengé, National Data Evaluation Center

Mounting statistical evidence continues to document an achievement gap between students along racial and economic lines. The gap is evident as early as kindergarten (West, Denton, & Reaney, 2000) on measures of letter recognition and letter-sound relationships between European Americans (Whites) and African-Americans and between Whites and Hispanic children. A similar gap is found along economic lines when the progress of poor children is compared to children who are not poor (West, Denton, & Germino-Hausken, 2000; Zill & West, 2000). This achievement gap, already evident at the beginning of kindergarten, can be found in first grade (Denton & West, 2002) and fourth grade (U.S. Department of Education, 2001). In fact, only Asians/Pacific Islanders have shown an improvement in their scores since 1992; African-Americans, Hispanics, and American Indians continue to

score at lower levels than Whites (U.S. Department of Education, 2001).

We wondered whether intervening early and providing specialized one-to-one teaching as soon as children begin to fall behind could make a difference to the achievement gap. Specifically, research already clearly shows that Reading Recovery, a one-to-one early literacy intervention, accelerates the progress of the lowest-performing children in first grade to reach average levels of performance (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994), but is it a complex enough response to make a difference in the achievement gap that exists along racial/ethnic and economic lines?

Theoretical Frame

Understanding the Achievement Gap

The concept of *closing the gap* stems from years of research that shows cer-

tain demographic groups academically underperform relative to other groups along racial and economic lines.

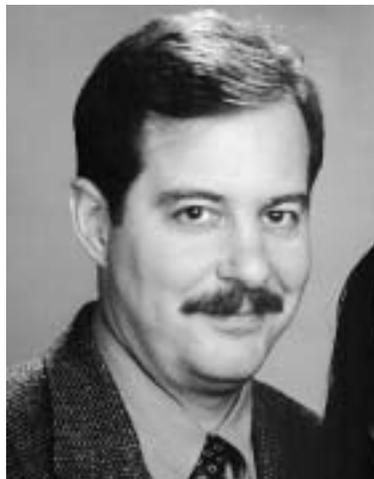
Achievement gaps are not limited to the United States; similar gaps have been documented between minority and majority cultures in other countries. In New Zealand for example, Maori and Pacific Island children typically achieve at lower levels than other children, and children in less economically advantaged schools achieve at lower levels than children in schools that are more economically advantaged (McNaughton, Phillips, & MacDonald, 2003).

It is not known why an achievement gap exists, but race itself is certainly not a determining factor. It is far more likely a result of an interaction of social, familial, and economic factors (Bainbridge & Lasley, 2002). Research shows that parents' education level and the economic level of the family are more meaningful predictors of school achievement than race (Bainbridge & Lasley, 2002, p. 424). These interacting factors mean that individual children come to school with different experiences, making some children less prepared for school than others. As Bainbridge and Lasley note, "Prior learning influences future achievement for all students, regardless of race" (p. 427).

A second set of factors that appears to contribute to achievement differences, in addition to those that exist at the level of the individual, are those at the school level (Land & Legters, 2002).



Emily M. Rodgers



Francisco Gómez-Bellengé

School climate, in terms of expectations for student progress, and policies such as tracking and retention all have a negative impact on student achievement and are likely to represent differences that become manifested as an achievement gap along racial and economic lines (p. 15).

What is clear is that the achievement gap must be closed and that it will take more than a quick fix, such as mandating a phonics program or emphasizing direct teaching, to compensate for the differences that exist (Bainbridge & Lasley, 2002). Nor can the problem be left up to individual teachers to solve, because as Bainbridge and Lasley note, change is needed on a grander scale to compensate for the inequities present in our society.

Data from the Ohio Fourth-Grade Proficiency Test reflect a similar literacy achievement gap exists in Ohio. Depending on the measure, Ohio educators are facing achievement gap differences of between 27 and 29 percentage points in fourth-grade reading performance when comparisons are made across racial/ethnic or economic status lines. In fact, the state's annual report card for 2002 contains the disturbing observation, "The race and ethnicity gap is clear and consistent" (Ohio Department of Education, 2002).

Early Intervention and One-To-One Teaching

A substantial body of research demonstrates that students who are having difficulty learning to read early in their academic career continue to struggle in later years (Juel, 1988; Vellutino & Scanlon, 2002). Since it is possible to use measures of literacy progress to identify children who are

having extreme difficulty learning to read even after just one year of school instruction (Clay, 2002), it seems critical to provide these children with an early intervention in order to spoil the prediction of failure that comes with early difficulties. One-to-one teaching is a recognized form of intervention and is becoming more widely used to prevent literacy failure (Wasik & Slavin, 1993). In particular, Reading Recovery, with its emphasis on teacher professional development, has been shown to be very effective in accelerating the progress of children having great difficulty learning to read (Shanahan & Barr, 1995).

Reading Recovery, however, has never been evaluated to determine whether students who receive the intervention make differential progress along racial and economic lines. In other words, do children who receive the intervention close the achievement gap between them and their classmates?

To examine this problem, we selected Reading Recovery as the early intervention context and we posed the following research questions:

1. Does a literacy achievement gap exist within a comparison group of first-grade students on fall-spring scores
 - a. across race/ethnicity lines?
 - b. across school lunch status lines (our proxy for economic status)?
2. For all students receiving a full Reading Recovery treatment (having an opportunity for 20 weeks of lessons), does the fall-spring gap in reading achievement relative to the comparison group decrease after Reading Recovery
 - a. across race/ethnicity lines?
 - b. across school lunch status lines?
3. For only those students who were successfully discontinued (successfully completed the intervention), does the fall-spring gap in reading achievement decrease
 - a. across race/ethnicity lines?
 - b. across school lunch status lines?

Methodology

All students in the study completed the six tasks published in *An Observation Survey of Early Literacy Achievement* (Clay, 2002) at two points in time: the first and last months of the 2001–2002 school year. We first disaggregated results for a comparison group of first-grade students in Ohio schools with Reading Recovery to determine whether or not a literacy achievement gap existed along race/ethnicity and economic lines. We then disaggregated results for Reading Recovery students along the same lines and compared their progress to that of the comparison group. We wondered if the African-American students who had Reading Recovery closed the gap between the White and African-American students in the comparison group. Similarly, we wondered whether less economically advantaged Reading Recovery students would perform more like poorer students in the comparison group or close the achievement gap with their more economically advantaged peers.

The sections that follow provide demographic information about the students and the Observation Survey, the assessment tool used to chart their progress over first grade.

Participants

Data were collected on 7,282 students who received Reading Recovery in Ohio during the 2001–2002 school year and on a comparison group of 1,915 first-grade students who did not receive Reading Recovery.

Comparison Group. Reading Recovery teachers randomly selected 100 students from the first-grade population at each of the 19 Reading Recovery sites in Ohio (sites are made up of several school districts). This population excludes students whose Observation Survey scores in the fall were low enough for them to qualify for Reading Recovery. Observation Survey data were collected by the Reading Recovery teachers at the beginning and end of the school year for the comparison group and were aggregated into a single state comparison group for this study ($n=1,915$). Because this group was selected from the population of first graders who did not qualify for Reading Recovery, they represent approximately the top 80% readers in their respective schools, with Reading Recovery students representing the lower 20% in those same schools in fall of first grade.

Reading Recovery Treatment Group. A total of 7,234 students received Reading Recovery lessons in Ohio in 2001–2002; of those, 5,547 had an opportunity to receive a full treatment, or a full series of lessons, defined as 20 weeks of instruction, whether they successfully completed the intervention or not. Of the remaining children who received Reading Recovery lessons, 1,687 children did not have an opportunity to receive a full series of lessons so they were excluded from the treatment group. Among those not included

were children who (a) moved while being served ($n=298$); (b) were still in Reading Recovery when the school year ended and ran out of time to have a full series of lessons ($n=1,197$); or (c) were removed from the intervention for unusual reasons, such as being placed back in kindergarten ($n=192$).

Discontinued Group. The discontinued group is a subset of the full treatment group. These are children whose programs were discontinued because they successfully met the rigorous criteria to complete the intervention. Of the 5,547 children who had an opportunity to receive a full intervention in Ohio during the 2001–2002 school year, 4,206 students' programs were successfully discontinued.

Disaggregated Lines

These three groups—the comparison group, treatment group, and the discontinued group—were disaggregated along race/ethnicity and economic lines in the following ways.

Race/Ethnicity. Each student's race/ethnicity was obtained from their school records when Observation Survey data were collected. Only two groups are examined in this study: children identified as White ($n=4,453$ Reading Recovery students and $n=1,223$ comparison group students) and children identified as Black/African-American, referred to in this study as African-American. ($n=853$ Reading Recovery students and $n=191$ comparison group students). We collected data on but did not examine results for other groups (Hispanic, Asian/Pacific Islander) because their numbers were too small to yield meaningful comparisons.

School Lunch Status. School lunch status was used as a proxy measure for

poverty or economic status. Data were taken from students' school records. Because this information cannot be required, data were not available on all children in the comparison, treatment, or discontinued groups. Students are classified as either

- regular lunch: These are children who did not receive any subsidy for their school lunches, either because their parents/guardians did not apply for the subsidy or the family did not meet federal criteria for poverty ($n=2,311$ Reading Recovery students and $n=753$ comparison group students).
- free lunch: These students received fully subsidized school lunches because their families were below the federal poverty line ($n=1,643$ Reading Recovery students and $n=276$ comparison group students).

Some students received partially subsidized lunches; however, their numbers were small and they were excluded from this analysis.

Test Measures

The Observation Survey is a group of standard measures developed in research studies with established reliabilities and validities indices (see Clay, 1993 or 2002). Stanines for Ohio students are reported for three points in time (fall, mid-year, and spring) for each task in Clay, 2002. Data were gathered and analyzed on all six Observation Survey measures, but only two measures, Text Reading Level and Hearing and Recording Sounds in Words, are reported here. These two measures are discussed here because they are regarded as valid measures of reading and writing progress (see for example,

Table 1. Fall–Spring Raw Scores and Stanines for the Comparison Group on the Hearing and Recording Sounds in Words (HRSW) and Text Reading Level (TRL) Measures Disaggregated by Race/Ethnicity

	Fall		Spring	
	Raw Score	Stanine	Raw Score	Stanine
HRSW – White	24.2	5	35.7	7
HRSW – African-American	21.9	5	34.7	6
TRL – White	5.2	7	21.4	7
TRL – African-American	3.8	7	17.6	6

McNaughton, Phillips, & MacDonald, 2003, p. 711). Results for the other four measures, including Letter Identification, Ohio Word Test, Concepts About Print, and Writing Vocabulary, did not disconfirm the trends reported here, and therefore we considered it redundant to discuss them.

Text Reading Level. Running records are used to obtain the instructional text reading level for individual students and to collect evidence of a child's independent strategic reading activity. The child reads aloud texts that represent a gradient of difficulty, until the highest text level with 90% accuracy or better is reached. The tester records text reading behaviors as the child reads.

Standard passages are used which are drawn from established basal systems

and have, over 20 years, proven to be a stable measure of reading performance. The range of scores and their basal reader grade level equivalents are as follows: Text Levels 0–2 = readiness; Levels 3–8 = preprimer; Levels 9–12 = primer; Levels 14–16 = end of Grade 1; Levels 18–20 = Grade 2; Levels 22–24 = Grade 3; Levels 26–30 = Grades 4–6.

Hearing and Recording Sounds in Words. Hearing and Recording Sounds in Words is a measure of phonemic awareness. In this task, the student writes a dictated sentence containing 37 phonemes. Each recorded phoneme is worth one point, yielding a possible score from 0–37. This task contains a closed set of 37 phonemes, representing the set of phonemes that children might be expected to be able to record. This

phonemic awareness measure has high validity and reliability (Clay, 2002).

Analysis

Two types of data are reported and analyzed. Raw scores from the Text Reading Level and Hearing and Recording Sounds in Words measures are reported first as group averages. These raw scores were then converted to stanines. Stanines for each measure were recorded for beginning and end of year for the following groups:

- comparison group of non-Reading Recovery students
- Reading Recovery Full Treatment
- Reading Recovery Discontinued

The groups were disaggregated by race/ethnicity and by categories of lunch cost. Stanines were plotted for the group average scores at fall and end of year. Comparisons were made across the groups.

Results

In this section, we address each research question in turn. Results are reported by measure and by disaggregated group.

Research Question #1: Does a literacy achievement gap exist within a comparison group of first-grade students?

Progress of the comparison group, disaggregated by race/ethnicity, on the Hearing and Recording Sounds in Words measure. When the comparison group of non-Reading Recovery students was disaggregated by race and results compared, we found a gap in raw scores in the fall on the Hearing and Recording Sounds in Words measure but no difference in stanines between the two groups. By spring, however, a gap in their sta-

Table 2. Fall–Spring Raw Scores and Stanines for the Comparison Group on the Hearing and Recording Sounds in Words (HRSW) and Text Reading Level (TRL) Measures Disaggregated by Economic Status

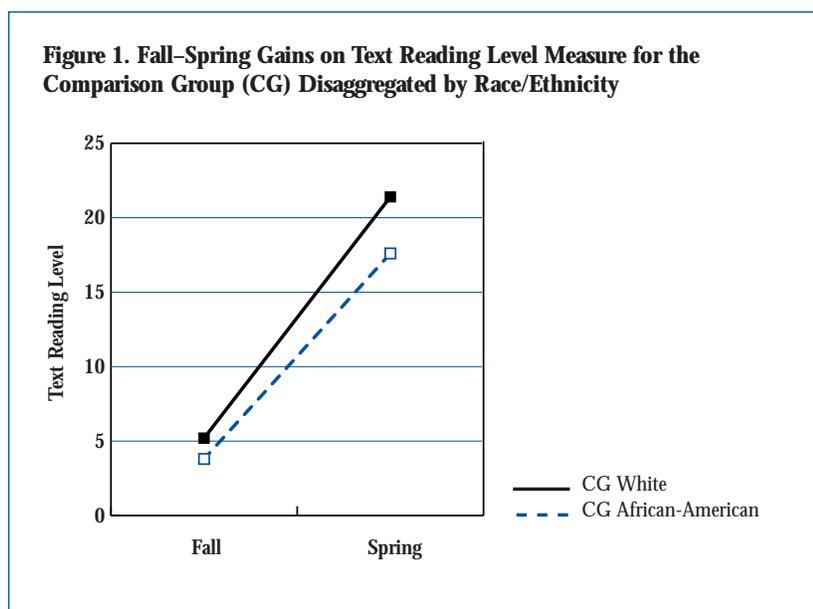
	Fall		Spring	
	Raw Score	Stanine	Raw Score	Stanine
HRSW – Regular Lunch	25.6	6	36.0	7
HRSW – Free Lunch	20.3	4	34.4	5
TRL – Regular Lunch	5.9	8	22.1	7
TRL – Free Lunch	3.3	6	16.9	6

nines emerged (see Table 1). Even though African-American and White children in the comparison group started first grade at a similar level on Hearing and Recording Sounds, they ended the year with a gap in their achievement.

Progress of the comparison group, disaggregated by race/ethnicity, on Text Reading Level measure. A similar trend was observed on the Text Reading Level measure. African-American children and the White children in the comparison group scored at the same stanine level in the fall but by spring, a gap emerged between the two, and African-American students finished at a lower level than the White students (see Table 1).

Progress of the comparison group, disaggregated by economic status, on Hearing and Recording Sounds in Words measure. Differences in the Hearing and Recording Sounds in Words measure were also evident when data were disaggregated on economic lines. On this measure, a gap of 2 stanines was present in the fall and in the spring between comparison group children eligible for free school lunch and those receiving regular-priced lunch. In this case, the gap was present in the fall and remained open through year-end (see Table 2).

Progress of the comparison group, disaggregated by economic status, on Text Reading Level measure. Table 2 also shows a gap along economic lines on the Text Reading Level measure, with a gap of 2 stanines in fall of first grade and 1 stanine in spring. On this measure, evidence of a slight closing of the gap was observed, according to stanine scores between comparison group students who qualify for free



lunch and those on regular-priced lunch (see Table 2).

Summary of Results for Research Question #1

To summarize, we first asked whether a gap was evident in the comparison group. We used two measures (Text Reading Level and Hearing and Recording Sounds in Words) and disaggregated the data two ways: along race/ethnicity and economic lines.

We found evidence of an achievement gap along both lines within the comparison group, reflecting a trend found in national and state data. Gaps widened on both measures along race/ethnicity lines which can be seen when raw scores are graphed (see Figure 1 for example).

When data were analyzed along economic lines in the comparison group, we found that the gap persisted for the Hearing and Recording Sounds in Words measure, and even though the gap on Text Reading Level appears to be closing according to stanine scores (shown in Table 2), when raw scores

for Text Reading Level were plotted, an opening gap became apparent. In other words, these gaps either remained or widened through first grade.

Research Question #2: For all students receiving a full treatment (having an opportunity for 20 weeks of lessons), does the fall-spring gap in reading achievement relative to the comparison group decrease after the Reading Recovery intervention?

Having established that an achievement gap exists along racial and economic lines in the comparison group, we examined the progress of all students who received a full series of Reading Recovery lessons, including those who were successfully discontinued and those who were not. We report first the outcomes of the treatment group, and then in response to Question #3, we consider only those students who were successfully discontinued.

Table 3. Fall–Spring Raw Scores and Stanines for the Comparison (CG) and Reading Recovery Treatment (RRTG) Groups on the Hearing and Recording Sounds in Words (HRSW) and Text Reading Level (TRL) Measures Disaggregated by Race/Ethnicity

	Fall		Spring	
	Raw Score	Stanine	Raw Score	Stanine
HRSW – CG White	24.2	5	35.7	7
HRSW – RRTG African-American	12.1	3	34.5	6
TRL – CG White	5.2	7	21.4	7
TRL – RRTG African-American	0.8	4	16.0	6

Progress of full treatment students, disaggregated by race/ethnicity, on Hearing and Recording Sounds in Words measure. In fall of first grade, on the Hearing and Recording Sounds in Words measure, African-American Reading Recovery students identified 12 sounds while White comparison group students identified 24; this difference corresponded to a gap of 2 stanines. By spring, the treatment group identified on average one fewer sound, with a gap of 1 stanine still evident (see Table 3).

The African-American students who

had an opportunity to receive a full series of Reading Recovery lessons were not able to completely close the gap with the White students in the comparison group, but they did reduce it. This trend is the opposite of that observed in the comparison group population in which the gap between African-American and White students widened on Hearing and Recording Sounds in Words over first grade (see Table 3).

Progress of full treatment students, disaggregated by race/ethnicity, on Text Reading Level measure. The fall

gap between the White students in the comparison group and the African-American students who had had an opportunity to receive a full series of Reading Recovery lessons was also wide on the Text Reading Level measure. The White students averaged a 5.2 text level in the fall and African-American treatment group students a 0.8 level, a difference equivalent to 3 stanines. By spring, the difference decreased to 1 stanine, but the African-American Reading Recovery students were still about 5 text levels below the Whites in the comparison group (see Table 3).

Progress of full treatment students, disaggregated by economic status, on Hearing and Recording Sounds in Words measure. In the fall, Reading Recovery students eligible for free school lunch identified 14 fewer sounds than the comparison group children who received regular lunch, a difference of 3 stanines. In spring, the difference in mean scores was fewer than 2 points, but the score still accounted for a gap of 2 stanines (see Table 4).

Progress of full treatment students, disaggregated by economic status, on Text Reading Level measure. On the text reading measure, the difference between the treatment group students on free lunch and comparison group students who qualified for regular priced lunch was 5 text levels or 4 stanines in the fall. This gap in stanines narrowed to 1 stanine in the spring (see Table 4).

Table 4. Fall–Spring Raw Scores and Stanines for the Comparison (CG) and Reading Recovery Treatment (RRTG) Groups on the Hearing and Recording Sounds in Words (HRSW) and Text Reading Level (TRL) Measures Disaggregated by Economic Status

	Fall		Spring	
	Raw Score	Stanine	Raw Score	Stanine
HRSW – CG Regular Lunch	25.6	6	36.0	7
HRSW – RRTG Free Lunch	11.3	3	34.3	5
TRL – CG Regular Lunch	5.9	8	22.1	7
TRL – RRTG Free Lunch	0.8	4	15.2	6

Summary of Results for Research Question #2

With Research Question #2, we compared outcomes for the treatment group (all Reading Recovery students

who had an opportunity to receive a full series of lessons, whether they were successfully discontinued or not) and the comparison group. While large gaps were evident along race/ethnicity and economic lines in fall of first grade between the treatment group and the comparison group, these gaps were narrowed by spring on both literacy measures—a contradictory trend to the one seen within the comparison group where gaps widened.

In the next section we look at a subset of the Reading Recovery treatment group: only those students who were successfully discontinued.

Research Question 3: For those students who were successfully discontinued, does the fall-spring gap in reading achievement decrease along race/ethnicity lines and along school lunch status lines?

In 2001–2002 in Ohio, 76% of those students who received a full series of lessons discontinued successfully from the intervention (n=4,206; Rodgers, Gómez-Bellengé, & Fullerton, 2003). We disaggregated the results for the discontinued and the comparison groups and compared their progress on the same two measures of the Observation Survey.

Progress of discontinued students, disaggregated by race/ethnicity, on Hearing and Recording Sounds in Words measure. African-American Reading Recovery students who would eventually discontinue from the intervention identified 10 fewer sounds than White students in the comparison group in the fall. By spring, this gap completely closed, both in terms of raw scores and stanines (see Table 5).

Table 5. Fall–Spring Raw Scores and Stanines for the Comparison (CG) and Discontinued (RRDisc) Groups on the Hearing and Recording Sounds in Words (HRSW) and Text Reading Level (TRL) Measures Disaggregated by Race/Ethnicity

	Fall		Spring	
	Raw Score	Stanine	Raw Score	Stanine
HRSW – CG White	24.2	5	35.7	7
HRSW – RRDisc African-American	13.8	3	35.6	7
TRL – CG White	5.2	7	21.4	7
TRL – RRDisc African-American	0.9	4	18.4	6

Progress of discontinued students, disaggregated by race/ethnicity, on Text Reading Level measure. A large gap of 4 text levels or 3 stanines was evident in the fall on the Text Reading Level measure between the African-American students who would be discontinued and comparison group Whites. By spring, the gap narrowed, but did not close, to 3 text levels or 1 stanine (see Table 5).

Progress of discontinued students, disaggregated by economic status, on Hearing and Recording Sounds in Words measure. Our final set of com-

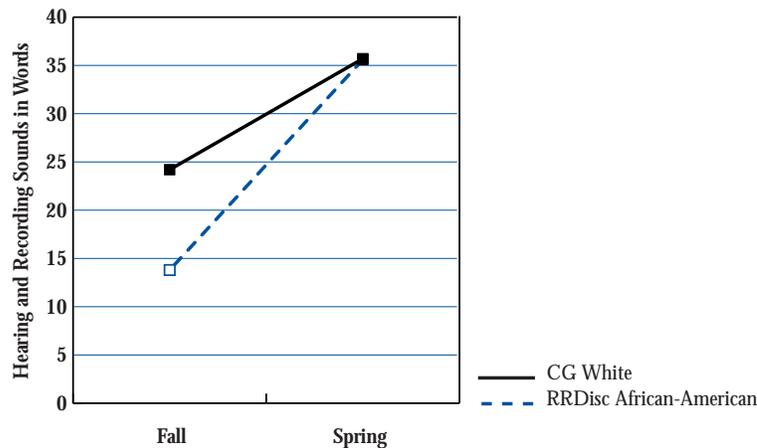
parisons is between the discontinued children who qualified for free lunch and children in the comparison group who received regular-priced lunch. For the Hearing and Recording Sounds in Words measure, a gap of 3 stanines was evident in fall. By the end of first grade, their raw scores and stanines were identical, and the gap had closed (see Table 6).

Progress of discontinued students, disaggregated by economic status, on Text Reading Level measure. A large gap of 5 text levels and 4 stanines existed in fall between the discontin-

Table 6. Fall–Spring Raw Scores and Stanines for the Comparison (CG) and Discontinued (RRDisc) Groups on the Hearing and Recording Sounds in Words (HRSW) and Text Reading Level (TRL) Measures Disaggregated by Economic Status

	Fall		Spring	
	Raw Score	Stanine	Raw Score	Stanine
HRSW – CG Regular Lunch	25.6	6	36.0	7
HRSW – RRDisc Free Lunch	13.1	3	35.6	7
TRL – CG Regular Lunch	5.9	8	22.1	7
TRL – RRDisc Free Lunch	0.9	4	18.1	6

Figure 2. Fall–Spring Gains for the Discontinued Group (RRDisc) on Hearing and Recording Sounds in Words (HRSW) Measure Disaggregated by Race/Ethnicity



ued and comparison group students. In spring, the gap was 4 text levels but only 1 stanine (see Table 6).

Summary of Results for Research Question #3

Students who were discontinued from Reading Recovery closed the achievement gap along race/ethnicity and economic lines with their counterparts in the comparison group on the Hearing and Recording Sounds in Words measure. This closing of the gap can be seen when raw scores are plotted (see Figures 2 and 3).

On the Text Reading Level measure, a wide gap existed along both lines in the fall. This gap was narrowed, but not completely closed, by the spring.

This narrowing trend is the reverse of that seen along the same lines in the comparison group in which the achievement gap widened.

Discussion

Trends in the Findings

Three different trends were evident

across the groups examined. First, within the comparison group, a widening gap was apparent on both measures when the data were disaggregated along race/ethnicity and economic lines. This trend mirrors the disturbing one found in national and state level data and discussed in the introduction to this paper.

Second, the group of all students who received a full series of Reading Recovery lessons (the treatment group) narrowed, but didn't close, the achievement gap along race/ethnicity and economic lines with their counterparts in the comparison group. Perhaps not surprisingly, the gap didn't entirely close because the treatment group includes all students who received Reading Recovery, including those who did not successfully discontinue. Even so, narrowing the gap is significant because it is a reversal of the one seen in national data and in the comparison group of this study where the gap is widening. Narrowing the gap is especially noteworthy considering the composition of the comparison group. This sample contains

the top 80% of first-grade students and excludes all Reading Recovery students. In effect, this exclusion raises the bar for the lowest-achieving students whose progress is compared to higher-achieving students.

Finally, when we considered only those Reading Recovery students who were successfully discontinued, we found that they closed the gap on the Hearing and Recording Sounds in Words measure, and they narrowed it on the Text Reading Level measure. Both trends are noteworthy because, again, they are the reversal of what we saw in the comparison group in which the gap widened or remained open over the year.

Lingering Gap in Text Reading Level

Despite the progress of the Reading Recovery students, it may be surprising to find that a gap remains along both racial/ethnic and economic lines on the Text Reading Level measure for successfully discontinued students. To put this finding in perspective, we should consider the raw scores and what they represent. The African-American students who were discontinued ended the year reading on average Text Level 18, while the White students in the comparison group read Text Level 21 or 22 on average. Books that are leveled at 18 are complex texts and represent the kinds of books that average first-grade students should be able to read at the end of first grade. In fact, Peterson (1991, p. 135) banded books from Levels 16–20 because of their similar text characteristics, which she described as

- elaborated episodes and events
- extended descriptions
- links to familiar stories
- literacy language

- unusual, challenging vocabulary
- illustrations provide low support

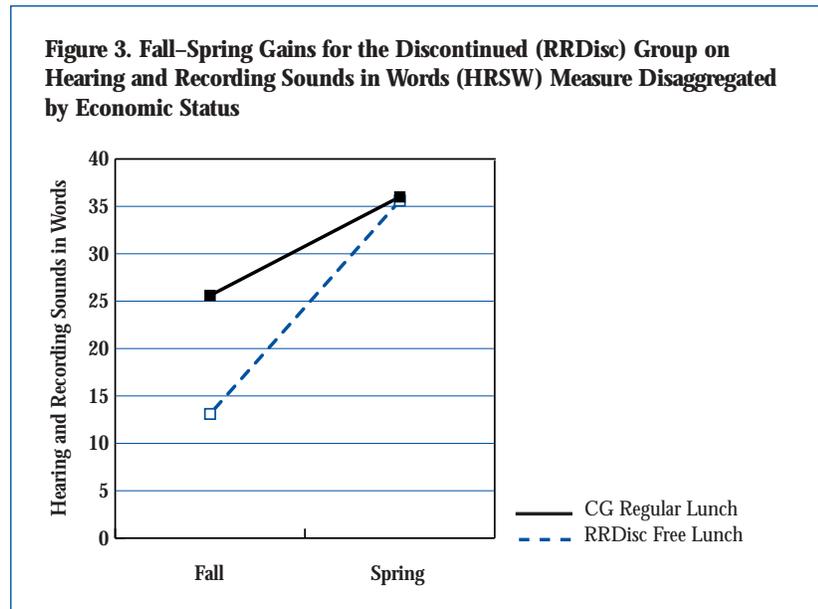
Hence, even though the gap didn't completely close for the text reading measure, we can state with some confidence that these discontinued students were reading texts that are, on average, the level of texts that most students would be expected to read at the end of first grade.

Closing the Achievement Gap

Bainbridge and Lasley (2002) make it clear that there are no quick fixes to the achievement gap. The interaction of individual as well as school-level factors is complex, and the problem will resist simple solutions. Educators can no more change a student's economic status than they can change the student's sex or parents' level of education. Magnifying the challenge is the fact that, while examples of particular schools closing the gap do exist (see Taylor & Pearson, 2002), success stories at a district or statewide level are rare if nonexistent (Skrla, Scheurich, & Johnson, 2000).

Reading Recovery does appear to make a difference to the gap in the achievement of different racial/ethnic and economic groups. Perhaps this is because it is a complex response to a complex problem. If differences in learning experiences prior to school have an impact on students' achievement and contribute to the achievement gap, it seems reasonable to expect that Reading Recovery, because it brings students up to an average level quickly before they fall too far behind, will have an impact on the achievement gap.

In addition to having an effect on factors at the individual level, Reading



Recovery also has an impact on school-level factors. Teachers receive professional development, more intensively at first while they are learning how to teach Reading Recovery and then ongoing professional development that provides continuing contact with colleagues (see Pinnell & Rodgers, in press, for a discussion of the importance of professional development). Professional development, delivered effectively, is related to improved student achievement (Rodgers & Pinnell, 2002). Reading Recovery also reduces retention rates and referrals to special education services (Gómez-Bellengé, 2002; Lyons, 1994), thereby having a further impact on school-level factors.

It appears from the findings of this research that there are added benefits to using Reading Recovery to deal with reading difficulties—that of closing or narrowing the literacy achievement gap that exists along economic and racial/ethnic lines. This is likely because Reading Recovery addresses multiple factors at the individual and school level that appear to be related to the achievement gap.

Conclusion

The state-level first-grade comparison group data reported here mirrors national and state findings of an achievement gap between African-Americans and Whites and between more and less economically advantaged students. Perhaps an even more disturbing finding was that on both the Hearing and Recording Sounds in Words and Text Reading Level measures, no achievement gap existed in the comparison group at the beginning of first grade; the children came to school without measurable differences evident on these two measures. It was after they had been in first grade for a year that the gap appeared.

We view this finding as particularly noteworthy because it underscores the need for educators and policy makers to investigate complex responses to a complex problem and points up the need for school districts to go beyond simple, short-term solutions. All evidence seems to suggest that Reading Recovery plays a significant role as part of a comprehensive literacy program, in not only bringing students

up to average reading levels, but also, at the same time, closing the achievement gap that develops in first grade along racial/ethnic and economic lines.

References

- Bainbridge, W. L., & Lasley, T. (2002). Demographics, diversity, and K-12 accountability: The challenge of closing the achievement gap. *Education and Urban Society, 34*, 422-437.
- Clay, M. M. (2002). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- Denton, K., & West, J. (2002). *Children's reading and mathematics achievement in kindergarten and first grade* (ERIC Document Reproduction Service No. ED461438).
- Gómez-Bellengé, F. X. (2002). External validation shows that Reading Recovery works. *Journal of Reading Recovery, 2*(1), 55-56.
- Juel, C. (1988). Learning to read and write: A longitudinal study of fifty-four children from first through fourth grades. *Journal of Educational Psychology, 80*(4), 437-447.
- Land, D., & Legters, N. (2002). The extent and consequences of risk in U.S. education. In S. Stringfield & D. Land (Eds.), *Educating at-risk students* (pp. 1-28). Chicago, IL: National Society for the Study of Education.
- Lyons, C. (1994). Reading Recovery and learning disabilities: Issues, challenges and implications. *Literacy, Teaching and Learning, 1*(1), 110-119.
- McNaughton, S., Phillips, G., & MacDonald, S. (2003). Profiling teaching and learning needs in beginning literacy instruction: The case of children in "low decile" schools in New Zealand. *Journal of Literacy Research, 35*, 703-770.
- Ohio Department of Education. (2002). *Committed to success for all: 2002 annual report on educational progress in Ohio*. Columbus: Author.
- Peterson, B. (1991). Selecting books for beginning readers. In D. DeFord, C. A. Lyons, & G. S. Pinnell (Eds.), *Bridges to literacy* (pp. 119-138). Portsmouth, NH: Heinemann.
- Pinnell, G., Lyons, C., DeFord, D., Bryk, A., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly, 29*, 8-39.
- Pinnell, G. S., & Rodgers, E. M. (In press). Reflective inquiry as a tool for professional development. In D. Strickland & M. Kamil (Eds.), *Improving reading achievement*. Norwood, MA: Christopher-Gordon Publishers.
- Rodgers, E. M., Gómez-Bellengé, F. X., & Fullerton, S. K. (2003). *Reading Recovery in Ohio: 2001-2002 State Report*. Columbus, OH: National Data Evaluation Center.
- Rodgers, E. M., & Pinnell, G. S. (2002). Professional development scenarios: What is and might be. In E. Rodgers & G. S. Pinnell (Eds.), *Learning from Teaching in Literacy Education: New Perspectives on Professional Development*. (pp. 1-9). Portsmouth, NH: Heinemann.
- Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly, 30*(4), 958-996.
- Skrla, L., Scheurich, J. J., & Johnson, J. (2000). *Equity-driven achievement-focused school districts: A report on systemic school success in four Texas school districts serving diverse populations*. University of Texas at Austin.
- Taylor, B., & Pearson, P. D. (Eds.). (2002). *Teaching reading: Effective schools, accomplished teachers*. Mahwah, NJ: Lawrence Erlbaum.
- U. S. Department of Education. (2001). *The nation's report card: Fourth-grade reading 2000*. Retrieved June 25, 2003, from <http://nces.ed.gov/nationsreportcard/data/ppt/1>
- Vellutino, F. R., & Scanlon, D. M. (2002). Research for the future: The interactive strategies approach to reading intervention. *Contemporary Educational Psychology, 27*, 573-635.
- Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28*, 179-200.
- West, J., Denton, K., & Germino-Hausken, E. (2000). Kindergartners' educational progress in reading: A study of language minority and non-language minority children. *Education Statistics Quarterly, 2*(1), 7-13.
- West, J., Denton, K., & Reaney, L. M. (2000). The kindergarten year: Findings from the early childhood longitudinal study, kindergarten class of 1998-99. *Education Statistics Quarterly, 2*(4), 25-30.
- Zill, N., & West, J. (2000). *Entering kindergarten: A portrait of American children when they begin school. Findings from the condition of education, 2000* (ERIC Document Reproduction Service No. ED448899).