ADDRESSING SUMMER READING SETBACK AMONG ECONOMICALLY DISADVANTAGED ELEMENTARY STUDENTS

RICHARD L. ALLINGTON and ANNE McGILL-FRANZEN
Department of Theory and Practice in Teacher Education, College of Education, Health, and Human Sciences, University of Tennessee, Knoxville, Tennessee

GREGORY CAMILLI
School of Education, University of Colorado, Boulder, Colorado

LUNETTA WILLIAMS
Department of Childhood Education, University of North Florida, Jacksonville, Florida

JENNIFER GRAFF
College of Education, University of Georgia, Athens, Georgia

JACQUELINE ZEIG
College of Education, University of Florida, Gainesville, Florida

COURTNEY ZMACH
American Institutes for Research, Washington, District of Columbia

RHONDA NOWAK
College of Education, University of Hawaii, Honolulu, Hawaii

Much research has established the contribution of summer reading setback to the reading achievement gap that is present between children from more and less economically advantaged families. Likewise, summer reading activity, or the lack of it, has been linked to summer setback. Finally, family socioeconomic status has been linked to the access children have to books in their homes and neighborhoods. Thus, in this longitudinal experimental study we tested the hypothesis that providing elementary school students from low-income families with a supply of self-selected trade books would ameliorate summer reading setback. Thus, 852 students from 17 high-poverty schools were randomly selected to receive a supply of self-selected trade books on the final day of school over a 3-year period, and 478 randomly selected students from these same schools received no books and served
as the control group. No further effort was provided in this intervention study. Outcomes on the state reading assessment indicated a statistically significant effect \( (p = .015) \) for providing access to books for summer reading along with a significant \( (d = .14) \) effect size. Slightly larger effects \( (d = .21) \) were found when comparing the achievement of the most economically disadvantaged students in the treatment and control groups.

A central aspect of federal educational policy making over the past half century has been attempts to fund interventions that close the achievement gaps that exist on measures of reading proficiency. The reading achievement gap between children from more and less economically advantaged families is substantial and has been persistent. Though there is evidence that this achievement gap has narrowed over time (Grissmer, Kirby, Berends, & Williamson, 1994), a recent National Assessment of Educational Progress in reading provides strong evidence of the pervasive nature of this seemingly intractable problem (Perie, Grigg, & Donahue, 2005). Though 77% of more economically advantaged fourth-grade students achieved above the basic level proficiency, only 46% of poor students (eligible for free or reduced-priced meals) achieved this modest level of reading development. The 27-point gap (203 vs. 230) is substantial and unchanged even with the advent of the latest federal initiative—the No Child Left Behind Act of 2001—targeted to closing the achievement gap (Gamse, Jacob, Horst, Boulay, & Unlu, 2009). But the failure of the various recent federal educational initiatives to close this reading achievement gap may stem from a failure of policy makers to focus attention on what seems a primary source of the existing achievement gap: summer reading setback.

Summer reading setback is a well-established phenomenon (Alexander, Entwisle, & Olson, 2007; Allington & McGill-Franzen, 2003; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Entwisle, Alexander, & Olson, 1997). Briefly, research on summer learning loss has provided reliable evidence that the reading achievement of economically disadvantaged students slides back a few months every summer. Cooper et al. (1996) reviewed 39 studies of summer academic loss and conducted a meta-analysis using 11 of the studies that provided sufficient data for that procedure. They found that “Middle-class students appeared to gain on grade-level equivalent reading recognition tests over summer
while lower-class students lost on them. There were no moderating effects for gender or race . . ." (p. 227). They concluded, “On average, summer vacations created a [reading] gap of about 3 months between middle- and lower-class students” (p. 261).

Although the role that summer reading setback plays in the reading achievement gap between economically disadvantaged and economically advantaged students was discussed some 20–30 years ago (Hayes & Grether, 1983; Heyns, 1978, 1987), that discussion produced little educational policy addressing either the issue or the impact. We designed this study as an initial step in developing an evidence base for interventions that might address summer reading setback and thereby narrow the persistent achievement gap between more and less economically advantaged children.

**Understanding Summer Setback**

Although there have been a number of studies establishing the problem of summer reading setback, there have been far fewer studies that have explored the reason that summer setback occurs. Additionally, not everyone finds a summer setback among poor children; rather, they report a lack of reading growth during the summer months among poor children while reading growth continues during the summer months for more advantaged children. Entwisle et al. (1997), for example, used a Fall to Spring assessment schedule and reported that more economically advantaged children added 47 raw score points over a 5-year period on reading achievement tests over summer vacations during elementary school years, whereas economically disadvantaged children added only 1 point. They developed the “faucet theory” to explain the phenomenon (Entwisle, Alexander, & Olson, 2001). In their view, when the school faucet is turned on—that is, when schools are in session—children of every economic background benefit roughly equally, but when the school faucet is turned off, as during summer vacations, reading proficiency among children from more economically advantaged families continues to develop, whereas no similar growth is observed in economically disadvantaged children.

On the other hand, Downey, von Hippel, and Broh (2004), analyzing the Early Childhood Longitudinal Study data, found
that economically disadvantaged children fell about 2.5 months behind more advantaged students during the summer months between kindergarten and Grade 1. The achievement gap grew far faster during the summer than during the school year. This finding parallels those reported in the Cooper et al. (1996) meta-analysis. But whether poor children’s reading proficiency decreases during the summer months or simply does not increase as much as more advantaged students seems largely irrelevant to the larger issue of the existing achievement gap. Irrelevant because in either case summer vacation periods seem to reliably produce differences in reading achievement among economically advantaged and disadvantaged children, small differences that expand over time such that Hayes and Grether (1983), using achievement data from the New York City public schools, estimated that as much as 80% of the reading achievement gap that existed between economically advantaged and disadvantaged students at sixth grade could be attributed to summer setback. Similar findings were reported more recently by Alexander et al. (2007) for ninth graders. In other words, each of these studies suggested that summer reading setback is a major contributor to the existing reading achievement gap between more and less economically advantaged children.

Summer Reading Activity

The research available points to summer reading activity, or the lack of it, as one source of summer reading setback. Heyns (1978) found that reading activity was the only factor that was consistently correlated to reading gains during the summer. She gathered a variety of data on children’s experiences during the summer vacation period but found that the number of books read, the amount of daily leisure reading, and the frequency of library use explained a larger proportion of the variation on a standardized test of word recognition than other recreational and enrichment summer activities. She concluded, “The single summer activity that is most strongly and consistently related to summer learning is reading” (p. 161).

However, given the evidence that economically disadvantaged children have limited access to books in the schools they attend and in their neighborhoods and homes (Allington, Guice,
Baker, Michelson, & Li, 1995; Constantino, 2005; Fryer & Levitt, 2002; Heyns, 1978; McGill-Franzen, Lanford, & Adams, 2002; Neuman, 1986; Neuman & Celano, 2001), various researchers have implicated this more restricted access to print as a primary source of documented differences in home reading activity by students from families at different levels of family income (Allington & McGill-Franzen, 2003; Cooper et al., 1996; Entwisle et al., 1997; Philips & Chin, 2004).

Research on Closing Summer Setback

There is a small set of studies focused on improving book access for children from low-income families during the summer months. These studies routinely report that simply supplying poor children with books during the summer months results in improved reading achievement. For instance, Kim (2004) conducted a post hoc analysis of a school district sponsored voluntary summer reading program for sixth-grade students and concluded that the effects of reading four or five books over the summer months on “on fall reading scores is potentially large enough to prevent a decline in reading achievement scores from spring to the fall” (p. 169). However, only a minority of the children actually read this many books, perhaps because the project did not supply the books to read. More recently, Kim (2006) reported that a single-year summer book distribution program had “marginally significant” effects on fall reading achievement. This marginally significant outcome may be the result of expecting too much from a single-year project. That is, even if providing books for summer reading improved student reading achievement as much as would be expected from attending school for 3 months, measuring that effect is difficult on most group achievement tests. Approximately 500 children from 10 elementary schools who had completed fourth grade were assigned to either receive eight books biweekly by mail or to a no book control group. The books were selected using a computer algorithm that identified books based on stated reading preferences and student lexile levels. Kim (2006) reported a marginally significant effect ($p = .059$) on reading achievement for students receiving the summer books with stronger effects ($p = .01$) observed for Black students than for Hispanic or White students. Finally, White and Kim (2008) reported on a more
complex single-year summer books intervention finding statistically significant effects on achievement when providing children from low-income families in Grades 3, 4, and 5 with free books and oral reading and comprehension scaffolding compared to children in the control group who received no summer books.

Thus, it seems that we now have initial evidence suggesting that ensuring access to books for summer reading can stem summer setback among economically disadvantaged children. But the effects observed to date have not been consistent or consistently significant. We believe that single-year interventions may simply provide insufficient support and that the effects of cumulative-year summer book interventions need to be tested.

In this study we tested the effects of one low-involvement intervention design in addressing summer reading setback. Our primary objective was to ensure that children from low-income families had easy access to books for voluntary summer reading over a 3-year period. Because self-selection of books has proven an important factor in fostering reading engagement (Guthrie & Humenick, 2004), we hypothesized that by providing these students with a collection of self-selected books at the end of each of 3 school years we could positively impact their voluntary summer reading activity and achievement and thereby ameliorate summer reading setback and, thus, close the reading achievement gap. We also designed our study to test the effects of access to self-selected books on beginning readers (Grades 1 and 2 at the start) because previous studies had targeted older students. Because we targeted younger students, we also supplied a larger number of books than previous studies. We did this because the lengths of books appropriate for beginning readers were invariably shorter than books appropriate for older readers.

**Method**

**Participants**

Students from 17 high-poverty elementary schools in two large school districts in Florida participated in the study. The enrollment of students eligible for free and reduced-price meals in these schools ranged from 65 to 98.5%. Most participating students were African American or Hispanic (89%), with few European
American participants (5%). When the study began the students were enrolled in either first or second grade. The age (rounded) distributions for each group of participants is given in Table 1.

All first- and second-grade students in these schools were given the parental consent letters to take home. We received 1,713 parental consent forms. Initially, 1,082 children were randomly selected to receive the summer books (treatment) and 631 children were randomly selected to serve as the no treatment (control) students. We selected substantially more treatment group students because of our concerns about participant attrition rates. For treatment students to remain in the study they had to be enrolled in one of the 17 schools where the book fairs were conducted. Thus, any participant who moved out of the 17 elementary schools could not continue as a treatment group member. However, the achievement of control students could be accessed from each districts’ student record files and so only the data for students who left the district would be lost. In the end, we had similar attrition for both groups. Attrition over the period of the study resulted in having data for 852 treatment (21.3% attrition rate) and 478 control (24.3% attrition rate) students. Much of the mobility in both groups of students was transfer between the 17 schools and so attrition rates from the experimental group sample were lower than we had anticipated originally.

For 3 consecutive years the children in the experimental treatment selected books during a book fair we ran to supply books for summer reading. The children in the control group received no trade books.1 The groups were deemed equivalent

---

**TABLE 1** Age Distribution of Participants at Year 3

<table>
<thead>
<tr>
<th>Age</th>
<th>Control</th>
<th>Treatment</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>123</td>
<td>213</td>
<td>336</td>
</tr>
<tr>
<td>10</td>
<td>228</td>
<td>400</td>
<td>628</td>
</tr>
<tr>
<td>11</td>
<td>110</td>
<td>204</td>
<td>314</td>
</tr>
<tr>
<td>12</td>
<td>17</td>
<td>32</td>
<td>49</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>478</td>
<td>852</td>
<td>1,330</td>
</tr>
</tbody>
</table>
on all relevant demographic characteristics at the end of the treatment period. There were no significant differences between students in the treatment and control groups on gender ($\chi^2 = 0.52, 1 df, p = .648$), free lunch status ($\chi^2 = 5.439, 3 df, p = .142$), race/ethnicity ($\chi^2 = 2.04, 4 df, p = .729$), age ($\chi^2 = 1.930, 5 df, p = .859$), or entering lexile level as assessed on the SRI ($t = -1.171, df = 963, p = .242$). The free lunch variable had four categories (free, reduced, not eligible, and not reported), and experimental treatment group children had a slightly higher frequency of free lunch eligibility at the beginning of the study.

**Dependent Variables**

We relied on the state mandated reading assessment as our primary measure of reading proficiency. The Florida Comprehensive Achievement Test (FCAT) is administered annually to all students in Grades 3 through 8. The FCAT has demonstrated reliability, with scores on the FCAT reading portion available as developmental scale scores with a range from 0 to about 3,000. Students should receive higher scores as they move from grade to grade according to their increased achievement (Florida Department of Education, 2001). We used the development scale scores in our analyses.

We elected to use the FCAT performances as our primary outcome measure because it is the high-stakes assessment that matters for both student promotion and school ratings. We compared the FCAT performances of participants in the treatment and control groups from the assessment administered in the year after 3 years of summer book distribution. In both groups there were students who were then enrolled in Grades 3, 4, and 5. Though normal school progress would have had all participants in Grades 4 or 5, the effects of retention in grade, in most cases mandated by state law (when Grade 3 FCAT performances fail to obtain the benchmark performance), resulted in some students being left back in one or more years.

**Participant Survey**

Survey data were gathered at the end of each summer from both treatment and control students using a shortened version of the
Literacy Habits Survey (Paris et al., 2004). This paper–pencil instrument was administered to students outside their classroom. The questions and possible responses were read aloud to the participants to ensure that they understood the questions being asked. Our primary interest was with items that requested information about summer reading activity, access to books, and home reading support, variables that earlier studies have suggested were potentially important in understanding voluntary summer reading activity/inactivity.

**Summer Book Distribution**

In the spring of each of the 3 years, we conducted a book fair at each school site. At the book fairs treatment students were brought in groups of 15 or so to the book fair area located in the school building. They were then given clipboards and pencils and were shown how to indicate which books they would like to receive (each book had a unique three-digit order number posted on the front cover). Students were told they could order 15 books and that they would receive 12 of those books to keep as their own (because of potential stock shortages we had students select more books than they would receive). Students then usually spent about 30 minutes reviewing the books available and indicating their selections on the order form. Once the order form was complete students returned to their classrooms. The books selected were individually boxed and delivered to each treatment student on the final day of school.

**Book Collections**

Each year the research team reviewed and selected a large number (400–600) of trade books for inclusion in the book fairs. The books were selected with a concern for text difficulty and interestingness. Each year we selected books in four broad categories.

**POP CULTURE**

These books featured characters and topics that were current in the broader popular media (television, movies, athletes, musicians, etc.).
SERIES BOOKS

Here we selected several titles from the most popular series books (as indicated by sales figures). Series books are not only popular with children but provide a number of benefits that make them useful for inexperienced or struggling readers (McGill-Franzen, 1993).

CULTURALLY RELEVANT

This category included books by minority authors and/or featured minority characters or regional themes.

CURRICULUM RELEVANT

Each year we reviewed the state science and social studies standards for topics that would be studied by the participants the following school year. We then selected titles that fit those topics.

Though each year we included titles reflecting each category, the numbers of books in each category varied over time. We attempted to avoid much overlap in titles each year. However, we did include some of the more popular titles from the previous year and worked to find sequels or additional books in a popular series or on a popular topic.

Overall, pop culture and series books were far more popular than book selections that were considered culturally relevant or curriculum related (Williams, 2008). In the final book fair 9 of the 10 most popular titles were from the pop culture or series book categories, and each of the 10 least commonly selected books were from the culturally or curriculum relevant categories.

Results

Reading Achievement

Descriptive statistics for comparing treatment and control groups are given in Table 2. Our first comparison tested the hypothesis that the FCAT performances of the treatment students would exceed those of the control group. A $t$-test found statistically significant differences ($t = 2.434$, $df = 1,328$, $p = .015$) in the performance of the treatment and control students on the FCAT administered after three consecutive summer book distributions. The effect size was also statistically significant ($ES = .14$).
Because there was a slight difference in the frequency of free lunch students between the treatment and control groups, these tests were also performed for students who were eligible for free lunch, excluding other socioeconomic status (SES) categories. Additional descriptive statistics (such as median and skewness) are given in Table 3 for free lunch students in the treatment and control groups. A \( t \)-test again found statistically significant differences \((t = 3.280, df = 1088, p = .001)\) in the performance of the free lunch–eligible students in the treatment and control groups on the FCAT administered after three consecutive summer book distributions. The effect size was slightly larger \((ES = .21)\). This indicates that the book distribution had positive effects on the reading achievement of the children from the lowest income families in the study.

### TABLE 2
Descriptive Statistics for the FCAT Developmental Scale Scores (DSS) in Year 3 (End of Treatment)

<table>
<thead>
<tr>
<th>Group</th>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Mean</td>
<td>1,238.34</td>
</tr>
<tr>
<td></td>
<td>(N)</td>
<td>478</td>
</tr>
<tr>
<td></td>
<td>(SD)</td>
<td>344.973</td>
</tr>
<tr>
<td>Treatment</td>
<td>Mean</td>
<td>1,285.56</td>
</tr>
<tr>
<td></td>
<td>(N)</td>
<td>852</td>
</tr>
<tr>
<td></td>
<td>(SD)</td>
<td>336.321</td>
</tr>
</tbody>
</table>

### TABLE 3
Robust Descriptive Statistics for Free Lunch–Eligible Students on the FCAT Developmental Scale Score at End of Treatment

<table>
<thead>
<tr>
<th>Group</th>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Mean</td>
<td>1,238.34</td>
</tr>
<tr>
<td></td>
<td>5% Trimmed mean</td>
<td>1,252.33</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>1,295.50</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>-.70</td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>344.97</td>
</tr>
<tr>
<td></td>
<td>IQR</td>
<td>422</td>
</tr>
<tr>
<td>Treatment</td>
<td>Mean</td>
<td>1,285.56</td>
</tr>
<tr>
<td></td>
<td>5% Trimmed mean</td>
<td>1,292.38</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>1,306.00</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>-.29</td>
</tr>
<tr>
<td></td>
<td>(SD)</td>
<td>336.32</td>
</tr>
<tr>
<td></td>
<td>IQR</td>
<td>425</td>
</tr>
</tbody>
</table>
At the end of the treatment, the Literacy Habits Survey (LH) was also examined to provide behavioral evidence of treatment implementation. In particular, treatment–control group differences were examined for two items: “How often did you read this summer?” (LHQ1: almost every day = 4, every week = 3, once or twice this summer = 2, not at all = 1); “Where do you get most of the books you read?” (LHQ6: school = 1, other = 0). The correlations of treatment membership with these implementation variables were \( r = .09 (p = .020) \) for LHQ1 and \( r = .28 (p < .001) \) for LHQ6. If these self-reported data are reliable, they suggest that our book distribution had a positive effect on the frequency of summer reading and that this resulted in the observed reading achievement gains in the experimental treatment subjects.

**DISCUSSION**

Summer reading setback has a demonstrated role in the current reading achievement gap between children from more and less economically advantaged families. It has been previously suggested that engagement in reading activity is an important factor in explaining summer setback (Heyns, 1978). Differential access to books has been offered as one explanation for the reported differences in voluntary reading activity between more and less advantaged students (Constantino, 2005; McGill-Franzen et al., 2002; Neuman, 1986; Neuman & Celano, 2001).

Our findings indicate that providing easy access to self-selected books for summer reading over successive years does, in fact, limit summer reading setback. We found that the experimental treatment group, which received the summer books for three consecutive summers, reported more often engaging in voluntary summer reading and had significantly higher reading achievement than the control group. We also found that the reading gains of students from the most economically disadvantaged families in the study were found to be larger, perhaps because these students have the most restricted access to books.

Our longitudinal outcomes are more consistent and larger than those reported for other summer book interventions (Kim, 2004, 2006; Kim & White, 2008). However, our study differs from earlier summer books studies in several ways:
Younger subjects completed Grades 1 and 2 at outset vs. completed Grades 3, 4, 5, or 6.

- Self-selected summer books vs. experimenter-selected books.

Each of these factors may be important in explaining the observed positive effects on reading achievement.

We assume that providing the books increased reading activity, as suggested here by the responses on the Literacy Habits Survey, where the treatment students regularly reported being more likely to receive books from their school (the summer books we distributed). The Literacy Habits Survey also indicated that treatment students were engaged in more reading activity during the summer months, suggesting that the summer book distribution did impact summer reading activity. Thus, this increased reading activity appeared to limit summer reading setback among the treatment group participants as indicated by the observed overall effect on reading achievement.

Though the effect sizes were not large ($ES = .14$ to $.21$), neither was the intervention effort or cost. Compare our outcomes with those reported by Cooper, Charleton, Valentine, and Mullenbruck (2000). They completed a meta-analysis of 54 studies and calculated that summer school attendance produced a positive effect on achievement that was identical ($d = .14$) to what our summer books intervention produced. Or consider the size of the effect (from $.09$ to $.14$) on achievement of adopting a school-wide reform model as calculated by Borman, Hewes, Overman, and Brown (2003). In other words, schools currently engage in reform and improvement efforts that are substantially more labor intensive and substantially more expensive but produce roughly the same effects on reading achievement as distributing free books for summer reading.

**Conclusions**

This study provides the best evidence to date that ensuring easy and continuing access to self-selected books for summer reading is one potential strategy for addressing summer reading setback and, therefore, addressing the reading achievement gap that
exists between students from more and less economically advantaged families. Future research should also include finer grained measures of summer reading activity in an attempt to more reliably link reading activity to improvements in reading achievement.

Theoretically, the self-teaching hypothesis developed by Share and Stanovich (1995) suggests one reason why voluntary reading, during the summer months or otherwise, would work to enhance reading development. According to the self-teaching hypothesis, each successful decoding encounter with an unfamiliar word provides an opportunity to acquire word-specific orthographic information. Such acquisition then influences reading automaticity and fluency, and, perhaps, comprehension and general reading development.

In addition, a variety of correlational and cross-sectional data (Allington, 1984, 2009; Anderson, Wilson, & Fielding, 1988; Cipielewski & Stanovich, 1992; Cunningham & Stanovich, 1998; Kim, 2004; Samuels & Wu, 2003; Stanovich, 2000; Taylor, Frye, & Maruyama, 1990) and experimental data (Kim & White, 2008; Kuhn, 2005; Kuhn et al., 2006) demonstrating links between reading volume and reading proficiency suggests one explanation for the effects of voluntary summer reading activity on reading development. We suspect that expanding reading activity, at least during the summer months, improves students’ reading achievements, if only because summer reading activity slows or reverses summer reading setback.

Roughly 30 years ago, Heyns (1978) suggested, “The unique contribution of reading to summer learning suggests that increasing access to books and encouraging reading may well have a substantial impact on achievement” (p. 172). Our data simply offer additional support for that conclusion.

Acknowledgments

This research was supported by grant # R305T010692-02 from the Office of Educational Research and Improvement, U.S. Department of Education. The two senior authors were coprincipal investigators for this longitudinal experimental study. The views expressed in this report reflect those of the authors.
Notes

1. Because this study was designed to experimentally test the impact of providing self-selected trade books, our design necessarily compared the impact of distributing self-selected books to some students and not others. We understood, as did the funding agency, our university Institutional Review Board, and the school districts’ personnel, that our design seemed to advantage students in the treatment group.

2. Our original design included an analysis based upon more assessment data but both school districts shifted their testing schemes dramatically as a result of demands by the state education agency during the implementation of the No Child Left Behind legislation. Rather than report the subskills assessment data that replaced reading assessment data, we turned to using the state reading assessment (Florida Comprehensive Achievement Test) administered beginning in Grade 3 because that remained the only constant reading assessment used in both districts.

References


