MATTHEW EFFECTS FOR WHOM?

Paul L. Morgan, George Farkas, and Jacob Hibel

Abstract. Which children are most at risk of experiencing a Matthew effect in reading? We investigated this question using population-based methodology. First, we identified children entering kindergarten on socio-demographic factors (i.e., gender, race/ethnicity, and socioeconomic status) known to index the relative risks and resources available to them as beginning readers. Second, we fitted growth curve models to the kindergarten-third-grade reading scores of these children as they participated in the Early Childhood Longitudinal Study-Kindergarten Class (ECLS-K) study. Third, we compared the relative reading achievement (as measured in standard deviation units from the sample’s overall mean across the study’s time points) of the children who were most and least at risk for reading disabilities. We found that the population subgroups most at risk for reading disabilities fell further behind typical readers over time. By contrast, those least at risk for reading disabilities did not move further ahead. Based on these findings, we conclude that a one-sided Matthew effect exists and that, moreover, it is likely to be experienced by children who are at greatest risk for reading disabilities.

The “Matthew effect” refers to a pattern of increasing advantage or disadvantage following initial advantage or disadvantage. The term comes from the Gospel according to Matthew: “For unto one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath” (XXV: 29, New Analytical). In reading, the “Matthew effect” (Stanovich, 1986) refers to the notion that “over time, better readers get even better, and poorer readers become relatively poorer” (Bast & Reitsma, 1998, p. 1373).

A specific developmental cycle, termed the Matthew effects model (Bast & Reitsma, 1998; Stanovich, 2000), is thought to result in the fan spread effect. That is, children in homes and schools fostering rapid acquisition of reading skills – the reading rich – should begin to enjoy reading from a very young age, and thus practice it more frequently. Frequent reading practice then fuels further skill acquisition (e.g., Cunningham & Stanovich, 1997; Guthrie, Schafer, & Huang, 2001). These children should spiral upward as increasingly competent and motivated readers. Hence, the “rich get richer.”

Children experiencing consistent difficulty acquiring reading skills – the reading poor – should follow a different trajectory. Their reading difficulties should lead them to develop more negative attitudes towards reading (e.g., Lepola, Salonen, & Vauras, 2000), and practice it far less (Anderson, Wilson, & Fielding, 1988; Scarborough & Parker, 2003). Over time, the “poor get...
poorer" because of this increasing avoidance of reading practice (Cunningham & Stanovich, 1997; Senechal, LeFevre, Hudson, & Lawson, 1996; Stanovich, 1986). Children with phonological deficits and other language or reading disabilities should be especially likely to experience this “poor get poorer” effect. Stanovich used the model to explain why children with learning disabilities can display increasingly generalized cognitive, motivational, and behavioral deficits.

Children with learning disabilities should be especially at risk of experiencing Matthew effects. This is not to say that learning disability is synonymous with reading disability (Lyon, 1996). Learning disabilities can occur in other skill areas, such as written expression or mathematical reasoning (Hallahan, Lloyd, Kauffman, Weiss, & Martinez, 2004). However, most children with learning disabilities have primary deficits in reading (Lyon, 1995). Snow, Burns, and Griffin (1998) estimated that about 80% of children with learning disabilities are poor readers. Stanovich (1986) put forth the Matthew effect theory as an etiological account of children’s identification as having learning disabilities.

Which groups of children are likely to experience the hypothesized Matthew effect in reading? We investigated this question using population-based methodology. First, we categorized children on the basis of their socioeconomic status (SES), gender, and race/ethnicity. Second, we used these factors to predict the intercepts and slopes in growth curve models of the children’s reading achievement from kindergarten through third grade. Third, we used the resulting estimated coefficients to calculate beginning and ending reading scores for subgroups of children categorized by differing combinations of SES, gender, and race/ethnicity. After standardizing the resulting scores using the standard deviations of reading proficiency in kindergarten and third grade (separately), we examined which, if any, of these population subgroups experienced Matthew effects in reading.

This methodology was not designed to focus on the detailed mechanisms by which Matthew effects may occur. Instead, we simply asked whether any of the population subgroups whose exogenous characteristics should predispose them to low or high reading performance indeed do experience a Matthew effect. Epidemiologically, the answer to this question is important in order to identify population subgroups that are most at risk of experiencing the model’s predicted generalized cognitive, motivational, and behavioral deficits. Identifying young children who are likely to lag increasingly behind in becoming readers should help researchers, practitioners, and policy-makers more effectively target early intervention services (McCoach, O’Connell, Reis, & Levitt, 2006; Parrila, Aunola, Leskinen, Nurmi, & Kirby, 2005). As detailed below, previous investigations of the Matthew effect have yielded inconsistent findings. A necessary first step in evaluating the validity of Stanovich’s (1986) theoretical model is to determine whether and for whom the predicted Matthew effect occurs.

Prior Studies of the Existence of Matthew Effects

Whether a Matthew effect truly occurs in reading has been much debated (e.g., Bast & Reitsma, 1997, 1998; Leppanen, Niemi, Aunola, & Nurmi, 2004; Parrila et al., 2005; Scarborough & Parker, 2003; Shaywitz et al., 1995; Stanovich, 2000). For such an effect to exist, two phenomena should be evident.

First, skill differences between good and poor readers should remain stable (i.e., “rich” readers remain rich whereas “poor” readers remain poor). Much empirical evidence exists for this phenomenon (e.g., Anderson et al., 1988; Baker, Decker, & DeFries, 1984; Jordan, Kaplan, & Hanich, 2002; Juel, 1988; McGee, Williams, Share, Anderson, & Silva, 1986; Scarborough, 1998; Shaywitz et al., 1995; but also see Phillips, Norris, Osmond, & Maynard, 2002). For example, Smith (1998) reported that, of children with the highest and lowest preschool assessment scores, 93% and 71% were reading above or below grade level, respectively, in the third grade. Juel found that 87% and 88%, respectively, of average-to-good and poor readers in first grade remained average-to-good and poor readers in fourth grade. Such stability led Juel to conclude that poor readers in her sample appeared “doomed” (p. 444).

Second, skill differences between good and poor readers should increase over time (i.e., rich readers become richer whereas poor readers become poorer). This increasing gap is sometimes referred to as the fan spread effect. The metaphor is apt, in that end points towards the base of an open fan are relatively close together but, traveling away from the base, they grow further apart. Much of the controversy surrounding the Matthew effect is due to inconsistent evidence for the fan spread effect (e.g., Leppanen et al., 2004). For example, whereas Bast and Reitsma (1997, 1998) reported that the gap in word recognition skills between good and poor readers widened over time, Aarnoutse and van Leeuwe (2000), Baker et al. (1984), and McCoach et al. (2006) did not find this to be the case.

Instead, poor readers have often been observed to narrow the reading achievement gap (e.g., Aarnoutse & van Leeuwe, 2000; Catts, Hogan, & Fey, 2003; Jordan et al., 2002; Parrila et al., 2005; Phillips et al., 2002; Shaywitz et al., 1995). This was the case in the only study to date to investigate the Matthew effect using a sample of children with learning disabilities.
Scarborough and Parker (2003) tracked a small sample (N = 57) of children with and without learning disabilities (i.e., reading disabilities, mathematics disabilities) from grade 2 to grade 8. Their analyses yielded a correlation of -.77 between the children's beginning reading scores and their reading growth. Given their own and the findings of others, Scarborough and Parker concluded that the Matthew effect remains "elusive, despite the plausibility and widespread acceptance of that well-reasoned hypothesis" (p. 65).

**Matthew Effects for Whom?**

Almost all previous studies have sought to test for a Matthew effect simply by identifying rich and poor readers as those with high and low reading test scores, respectively, at or near the beginning of their school careers and then comparing these scores to those from a later time point (e.g., Bast & Reitsma, 1998; Shaywitz et al., 1995). This approach is intuitively appealing, but it is problematic. For example, important information (e.g., gender, SES) about children's background characteristics is lost when they are classified simply as good or poor readers. Methodologically, children's performance (both at school entry and later) is at least partially due to chance elements and measurement error, as well as to circumstances that are subject to change.

This traditional approach does not explicitly account for the differing resources available to children from more or less advantaged families. For example, consider two children with the same initial low score in kindergarten. One may have a phonological processing deficit but receive a great deal of supplemental assistance over time from his highly educated and well-to-do parents. In contrast, the other child may have no such deficit, but receive no assistance from his poorly educated parents who are living in poverty. Which, if either, of these two children should be classified as a poor reader in a test of Matthew effects?

Additional factors will likely influence the Matthew effect's occurrence, and should thus be taken into account. One set of influences is made up of exogenous child- and family-level factors that help shape the context within which a young child's reading growth occurs. Examples include the child's gender, his or her race/ethnicity, and the family's social class background (e.g., D'Angiulli, Siegel, & Hertzman, 2004; McCooch et al., 2006; Neuman & Celano, 2001).

Another set of factors is the language- and literacy-related actions of a child or his/her parents or caregivers during the preschool period, as well as the language- and literacy-related resources available to each. Examples of such actions and resources include whether and to what extent the child (a) engages in shared storybook reading or visits the library, (b) accesses books at home, (c) converses with an adult who uses a relatively complex vocabulary, and (d) interacts with parents or caregivers who provide instruction in concepts about print and letter knowledge (e.g., Neuman, 1999; Snow, Barnes, Chandler, Goodman, & Hemphill 1991; Weigel, Martin, & Bennett, 2005). These variables affect the child's emergent literacy skills, such as phonological processing ability, knowledge of print concepts, emergent writing, oral vocabulary, and letter name and sound knowledge (e.g., Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Whitehurst & Lonigan, 2002).

A third set of influences is the reading-related actions undertaken by the child and his/her teachers, parents, and caregivers through the school years, as well as the reading-related resources available to the child. These variables include the curriculum and how a teacher chooses to deliver it, the child's peer group, the parental assistance with reading provided to the child, and, again, the child's own interest and reading efforts inside and outside school (e.g., Cunningham & Stanovich, 1997; Guthrie, Wigfield, Metsala, & Cox, 1999).

Consequently, testing for a Matthew effect is determined in part by which of these sets of diverse factors are included in the estimated model. Related to this choice is the conceptualization of which groups of children are considered likely to be reading rich and poor. This conceptualization is likely to strongly affect the results of any test for fan spread.

In this study, we defined rich and poor readers by exogenous child- and family-level background variables indexing the biological, social, and economic resources available to each child and his or her family. Thus, we asked, "Matthew effects for whom?"

This approach has both substantive and methodological advantages. First, it allows us to identify groups of children who, over subsequent years, will likely average stronger or weaker literacy-related abilities, interests, actions, and inputs (via parents, childcare workers, peers, and teachers). Such higher-versus lower-level flows of reading-related resources, activities, and instruction seem to us to constitute a particularly appropriate conceptualization of a child's reading-related "wealth."

Second, using exogenous child and family background variables allows us to better track the growth trajectories of children most at risk for reading disabilities. Epidemiological research has repeatedly found that children from certain population subgroups (i.e., boys, minorities, and those from low-income households) are much more likely to be identified as having disabilities (e.g., Delgado & Scott, 2006; Donovan & Cross, 2002; Kavale, 1988; Klinger, Artiles, & Barletta, 2006). For example, Katusic, Colligan, Barbaresi, Schaid, and Jacobsen (2001) noted that boys were two to three times
more likely to have reading disabilities than girls, regardless of whether a regression-, discrepancy-, or low-achievement identification method was used. Further, Artiles, Rueda, Salazar, and Higareda (2005) reported that children who were English Language Learners (ELL) were 3.5 times more likely to be placed in special education by 12th grade than children who were language proficient. Stanton-Chapman, Chapman, and Scott (2001) noted that low maternal education was the strongest child-level predictor of school-identified disability.

Moreover, interactions between these gender, race, and social class factors may further increase a child’s likelihood of having a disability. For example, whereas only about 7% of White mothers of school-aged children have less than a high school diploma, the comparable rates for African-American or Latino mothers are about 20% and 50%, respectively (NCES, 2002). Artiles et al. also reported that ELL children from low-income homes were 1.4 times more likely to be identified as having learning disabilities than ELL children from middle-to-high income homes. Because gender, race, and SES (a) are relatively well-established predictors of children’s relative reading skill levels and (b) can reasonably be characterized as exogenously related to the acquisition of such skill, we incorporated these child- and family-level variables into our statistical models. By doing so, we were able to more rigorously test for the occurrence of Matthew effects.

Third, this approach has the advantage that it can be empirically implemented by recently developed techniques of growth curve modeling (Goldstein, 1995; Raudenbush & Bryk 2002; Singer & Willett, 2003). Such models include a number of properties that are particularly useful in testing for the fan spread effect (Bast & Reitsma 1997; Campbell & Kenny 1999; Shaywitz et al., 1995). The estimated coefficients from growth curve models allow one to compute average reading starting values and growth trajectories for population subgroups defined by their background characteristics. These estimated coefficients and growth trajectories then reveal which groups of children are becoming stronger or weaker readers relative to typical children over time, empirically answering the question, “Matthew effects for whom?”

**METHOD**

*Design of the Study*

We tested for Matthew effects using a model that included both child- and family-level variables, as well as endogenous reading achievement outcomes as children progressed from the fall of kindergarten to the spring of third grade. In particular, we tested whether the reading trajectories of subgroups of children defined by gender, race/ethnicity, and parents’ SES demonstrated the Matthew effects property. That is, we asked: Do those most at risk for reading disabilities (i.e., boys, Blacks and Hispanics, and those arriving at school from low-income households) begin, on average, near the low end of the reading skills distribution and move further below the mean over time, whereas those least at risk for reading disabilities (e.g., girls, Asians, and those arriving from high-income families) begin near the top of the distribution and increase their advantage over time?

The answer to this question offers more than a simple “yes/no” test of whether children who began near the bottom or top of a particular reading skill measure’s distribution end up, respectively, further below or above the mean later on. Instead, it identifies which, if any, population groups of children defined by sociodemographic variables that index the reading-related resources available to them begin school as poor or rich readers and then experience systematic tendencies to grow relatively poorer or richer.

As noted, we focused on risk factors such as gender, race/ethnicity, and social class because many previous studies have found them to strongly predict the reading actions, resources, and skills growth of children, as well as their relative risk of reading disabilities (Kavale, 1988; McCoach et al., 2006). For example, the results of ethnographic (e.g., Lareau, 2003; Neuman & Celano, 2001), survey (e.g., Dickinson, McCabe, & Anastasopoulos, 2002), and quasi-experimental research (e.g., Dowhower & Beagle, 1998) all indicate that young children living in socioeconomically (SES) poor communities are particularly likely to begin school as poor readers, as they often lack access to books and other print materials. Further, less well-educated parents, caregivers, and child-care workers spend less time teaching knowledge of letters and letter-sound correspondence. They are also less likely to transmit the oral language skills (e.g., grammatical-syntactic coding and vocabulary knowledge) that are useful for the transition to school (Farkas & Beron, 2004; Hart & Risley, 1995; Heath, 1983; Whitehurst & Lonigan, 2002).

McCoach et al. (2006) recently reported that children from low-income families scored, on average, 6.2 point lower on a measure of reading proficiency (i.e., the Reading Test, the same measure used here) than children from high-income families across their first two years of school. Their analyses indicated that SES “is one of the most powerful predictors of both reading achievement at kindergarten entry and summer reading growth rate” (p. 25).

Racial and ethnic differences are also evident in the growth of children’s reading skills (e.g., Landgren, Kjellman, & Gillberg, 2003; Sanchez, Bledsoe, Sumabat,
& Ye, 2004), although, for some children, differences in the quality of education may at least partially explain these achievement discrepancies (e.g., Beron & Farkas 2004; Fryer & Levitt, 2004; Manly, Jacobs, Touradji, Small, & Stern, 2002). Gender also appears to moderate the effect of early reading struggles on children’s reading motivation and skill (e.g., Lepola, 2004; Riordan, 2002).

The ECLS-K Data
We estimated the model using data from the Early Childhood Longitudinal Study–Kindergarten Class (ECLS-K), a large, longitudinal, and nationally representative sample of U.S. schoolchildren who entered kindergarten in 1998 and whose reading progress is still being followed (Rathbun & West, 2004; Rock & Pollack, 2002; West, Denton, & Reaney, 2000). These data are collected and made available through the U.S. Department of Education’s National Center for Education Statistics (NCES). The database is a multistage cluster sample of elementary schools, classes within these schools, and children within these classes.

Schools were selected from geographic areas consisting of counties or groups of counties from which 1,280 public and private schools offering kindergarten programs were originally selected. A target sample of 24 children from each public school and 12 children from each private school was drawn, with Asian/Pacific Islander children oversampled. We analyzed data from 10,587 children across five time points (i.e., the fall and spring of kindergarten and first grade, and the spring of third grade).

Reading Test. The ECLS-K’s Reading Test was developed through a multi-step panel review process (see Rock & Pollack, 2002, for details). Items were included in the test’s final form if they displayed (a) acceptable item-level statistics, (b) good fit with maximum likelihood item response theory (IRT) parameters, and (c) no differential item functioning across gender or race (NCES, 2005).

The test consists of subtests of three main types of reading skills. The first skill category is basic skills, including familiarity with print and recognition of letters and phonemes. The second is vocabulary. The third is reading comprehension. Measures of reading comprehension were based on a National Assessment of Educational Progress framework involving four types of reading comprehension skills: (a) initial understanding; (b) developing interpretation; (c) personal reflection and response; and (d) demonstrating a critical stance (Rock & Pollack, 2002).

Utilizing one-to-one-administered adaptive testing, children were given a test whose coverage of these domains varied according to their grade and skill level (Rock & Pollack, 2002). Most of the Reading Test’s items utilize a multiple-choice format. A few are open-ended questions or call for a constructed response. The content emphasis changes over time as children grow as readers. For first graders, 40%, 10%, and 50% of the measure’s testing time is devoted to assessing basic skills, vocabulary, and comprehension, respectively. For third graders, these percentages change to 15%, 10%, and 75%, respectively.

The Reading Test displays very good psychometric properties. The ECLS-K data provide an overall Item-Response Theory (IRT) scale score, which serves as a composite summary measure of each child’s reading proficiency at each time point. The reliabilities of the IRT theta scores (the appropriate measure of internal consistency) on the full reading test range from .93 to .97 (NCES, 2000). First graders’ Reading Test scores correlated .85 or above with the Kaufman Test of Educational Achievement reading test (NCES, 2002); third graders’ scores correlated .83 with the Woodcock-McGrew-Werder Mini-Battery of Achievement (NCES, 2005).

Child and family characteristics. Child-level variables include the child’s gender, race/ethnicity (a set of dummy-coded variables comparing White children with Black children, Hispanic children of all races, Asian children, and children categorized as belonging to “other” races or ethnicities), and a standardized composite measure of the child’s family SES. The NCES-calculated SES variable is an average of each parent’s or guardian’s self-reported household income and each parent’s (or guardian’s) education, as well as each parent’s (or guardian’s) occupational prestige.

Data Analysis
We used multilevel linear growth curve modeling to analyze children’s initial level and growth rate in reading (Raudenbush & Bryk, 2002; Singer & Willett, 2003). The estimated model was specified as follows:

Level 1 equation:

\( y_{it} = b_{0i} + b_{1i}t + e_{it} \)

The intercept and slope regression coefficients from the first stage were then written as a function of exogenous background characteristics:

\( b_{0i} = c_0 + c_1X_{it} + \ldots + c_kX_{ki} + u_i \)
\( b_{1i} = d_0 + d_1X_{it} + \ldots + d_kX_{ki} + v_i \)

The “c” coefficients show the effects of exogenous background characteristics on starting values, whereas the “d” coefficients show the effects of background on the score growth rate. Thus, this model allows us to simultaneously identify the reading “rich and poor” (i.e., students with high and low estimated intercepts, respectively) and those who are growing “richer” or “poorer” (i.e., students with steep growth curves and
those with flatter curves, respectively). After estimating the coefficients in this model, we used the results to calculate predicted scores for the population subgroups that begin and end their reading performance trajectories at either the low or the high end of the reading score distributions.

RESULTS

Table 1 displays descriptive statistics for the analytical sample. The sample was 62% White, 13.5% Black, 13.2% Hispanic, 5.3% Asian, and 6% Other Race or Ethnicity (this included Native Americans and mixed-race children). With family SES coded into quintiles, the average score was 3.30. (This was above 3.0 because of a slight excess of missing cases below the mean of the variable.) The sample was 50% male.

The ECLS-K researchers computed a continuous IRT-scaled composite reading score for each child at each survey wave. As shown in Table 1, when kindergarten began, the mean of this composite was 22.91, with a standard deviation of 8.60. By the spring of third grade, both the mean and the standard deviation had increased substantially, to 109.44, and 19.45, respectively. These beginning and ending means and standard deviations can be used to compute, for poor and rich readers, whether their difference from the mean has or has not increased over time.

Table 2 presents the fitted growth curve coefficients for the composite reading score. The first column shows a simple regression with just a constant term and slope; the second column shows how these vary as a function of a child's characteristics. As expected, for the beginning score (y-intercept of the fitted growth curves), males performed lower than females, and SES had a strong positive effect. Hispanics and Other Race or Ethnicity performed lower than Whites; Asians performed higher. Blacks were not significantly different from Whites.

For the reading growth slope, males' reading skills grew more slowly than females'. SES had a significant positive effect on skills growth. Reading skill grew much

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive Statistics of the Study's ECLS-K Analytical Sample</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ECLS-K Reading Test</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRT Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall K</td>
<td>22.91</td>
<td>8.60</td>
<td>10.08</td>
<td>69.66</td>
</tr>
<tr>
<td>Spring K</td>
<td>32.26</td>
<td>10.43</td>
<td>10.85</td>
<td>70.80</td>
</tr>
<tr>
<td>Fall 1st</td>
<td>39.74</td>
<td>12.78</td>
<td>12.69</td>
<td>86.63</td>
</tr>
<tr>
<td>Spring 1st</td>
<td>57.09</td>
<td>13.26</td>
<td>14.77</td>
<td>88.95</td>
</tr>
<tr>
<td>Spring 3rd</td>
<td>109.44</td>
<td>19.45</td>
<td>42.36</td>
<td>148.95</td>
</tr>
<tr>
<td>White</td>
<td>0.62</td>
<td>0.49</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Black</td>
<td>0.14</td>
<td>0.34</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.13</td>
<td>0.34</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Asian</td>
<td>0.05</td>
<td>0.23</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Other Race or Ethnicity</td>
<td>0.06</td>
<td>0.24</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>SES Quintile</td>
<td>3.30</td>
<td>1.33</td>
<td>1.00</td>
<td>5.00</td>
</tr>
<tr>
<td>Gender (Male=1)</td>
<td>0.50</td>
<td>0.50</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note. N = 10,587; SD = Standard Deviation; Min = Minimum; Max = Maximum.
Table 2

Growth Curve Estimates for ECLS-K Reading Test IRT Score

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>15.58***</td>
<td>10.17 ***</td>
</tr>
<tr>
<td>Male</td>
<td>-1.37 ***</td>
<td>-1.88 ***</td>
</tr>
<tr>
<td>SES</td>
<td>1.88 ***</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.09</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-1.14 ***</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>3.36 ***</td>
<td></td>
</tr>
<tr>
<td>Other Race or Ethnicity</td>
<td>-0.83</td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td>2.15 ***</td>
<td>1.99 ***</td>
</tr>
<tr>
<td>Male</td>
<td>-0.05 ***</td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>0.07 ***</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.25 ***</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.06 ***</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>-0.07 ***</td>
<td></td>
</tr>
<tr>
<td>Other Race or Ethnicity</td>
<td>-0.17 ***</td>
<td></td>
</tr>
<tr>
<td>Tau</td>
<td>0.082</td>
<td>-0.039</td>
</tr>
</tbody>
</table>

Note. N = 10,587; *p < .05; ***p < .001.

more slowly for Blacks than for Whites. Children who were Hispanic, Asian, or Other Race or Ethnicity grew more slowly than children who were White.

Table 3 displays predicted beginning and ending Z-scores for population subgroups based on the growth curve coefficient estimates. These scores are computed for gender x race subgroups with SES at either the lowest or highest quintile. These are the subgroups with the lowest and highest average beginning and ending scores. They are calculated using the reading score means and standard deviations computed separately for kindergarteners and third graders. As indicated in Table 3, in the fall of kindergarten, the lowest performing groups were those in the lowest quintile of SES. Among these, the very lowest were Hispanic and Other Race or Ethnicity males, and Other Race or Ethnicity females.

These children performed .83-.77 standard deviations below the kindergarten mean. By contrast, the highest performing groups were those with SES in the highest quintile. Among these children, the very highest ranked subgroups were Asian females and males. These children performed 1.16-1.01 standard deviations above the mean.

By the spring of third grade, all subgroups were reading at a much higher level. However, relative to one another, generally, those in the lowest quintile of SES performed below the mean whereas those in the highest quintile performed above the mean. Nevertheless, several groups changed their relative ranking, as detailed below.

Black males in the lowest SES quintile lagged further behind their peers in reading growth. Between the
beginning of kindergarten and the end of third grade, their average reading Z-score declined from -0.66 to -1.12. That is, their reading skills fell an additional 0.5 of a standard deviation below that of typical readers. A similar result was observed for Other Race or Ethnicity group males, whose relative reading performance fell from -0.83 to -1.18. For Black females, reading performance fell from -0.63 to -0.93, a decline of 0.3 standard deviation. Other Race or Ethnicity females fell from -0.77 to -1.08, also a decline of about 0.3 standard deviations. Thus, the relative reading performance of these four groups of at-risk readers exhibited fan spread: Their relative reading performance became worse over time.

We observed no comparable increases in relative position among the groups of children who were initially the highest performing. Indeed, the relative positions of the groups who were the greatest distance above the mean when school began (i.e., Asian females and males) eroded by approximately 0.5 of a standard deviation by the spring of third grade. Furthermore, the largest group of high-performing children (i.e., high-SES White females) showed little change in their relative position.

| Table 3 |

| Standardized Beginning and Ending Scores on the ECLS-K IRT-Scaled Reading Test for Population Subgroups |

<table>
<thead>
<tr>
<th>ECLS-K Reading Test IRT Score</th>
<th>Beginning</th>
<th>Ending</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Male, SES 1</td>
<td>-0.48</td>
<td>-0.77</td>
</tr>
<tr>
<td>Black Male, SES 1</td>
<td>-0.66</td>
<td>-1.12</td>
</tr>
<tr>
<td>Hispanic Male, SES 1</td>
<td>-0.80</td>
<td>-0.77</td>
</tr>
<tr>
<td>Asian Male, SES 1</td>
<td>-0.40</td>
<td>-0.20</td>
</tr>
<tr>
<td>Other Male, SES 1</td>
<td>-0.83</td>
<td>-1.18</td>
</tr>
<tr>
<td>White Female, SES 1</td>
<td>-0.46</td>
<td>-0.46</td>
</tr>
<tr>
<td>Black Female, SES 1</td>
<td>-0.63</td>
<td>-0.93</td>
</tr>
<tr>
<td>Hispanic Female, SES 1</td>
<td>-0.65</td>
<td>-0.43</td>
</tr>
<tr>
<td>Asian Female, SES 1</td>
<td>-0.37</td>
<td>-0.42</td>
</tr>
<tr>
<td>Other Female, SES 1</td>
<td>-0.77</td>
<td>-1.08</td>
</tr>
<tr>
<td>White Male, SES 5</td>
<td>0.47</td>
<td>0.52</td>
</tr>
<tr>
<td>Black Male, SES 5</td>
<td>0.04</td>
<td>0.15</td>
</tr>
<tr>
<td>Hispanic Male, SES 5</td>
<td>0.23</td>
<td>0.28</td>
</tr>
<tr>
<td>Asian Male, SES 5</td>
<td>1.01</td>
<td>0.58</td>
</tr>
<tr>
<td>Other Male, SES 5</td>
<td>0.73</td>
<td>0.47</td>
</tr>
<tr>
<td>White Female, SES 5</td>
<td>0.62</td>
<td>0.68</td>
</tr>
<tr>
<td>Black Female, SES 5</td>
<td>0.33</td>
<td>0.06</td>
</tr>
<tr>
<td>Hispanic Female, SES 5</td>
<td>0.38</td>
<td>0.49</td>
</tr>
<tr>
<td>Asian Female, SES 5</td>
<td>1.16</td>
<td>0.67</td>
</tr>
<tr>
<td>Other Female, SES 5</td>
<td>0.58</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Note. SES reported in quintiles; 1 = lowest, 5 = highest.
Thus, we observed no fan spread for resource-rich, low-risk population subgroups. In other words, as far as literacy acquisition is concerned, a “one-sided Matthew effect” appeared to exist – the poor grow poorer, but the rich do not grow richer.

**DISCUSSION**

The existence of a Matthew effect in reading is controversial (e.g., Bast & Reitsma, 1997, 1998; McCoach et al., 2006; Parrila et al., 2005; Scarborough & Parker, 2003; Shaywitz et al., 1995; Stanovich, 2000). Much of the controversy is due to inconsistent evidence for the fan spread effect (e.g., Aarnoutse & van Leeuwe, 2000; Bast & Reitsma, 1998; Scarborough & Parker; Shaywitz et al.). Previous studies have tested for the fan spread effect by comparing children’s performance on a reading measure near the beginning of their school careers to scores from a later point (e.g., Bast & Reitsma, 1998; Shaywitz et al.). The methodological and substantive limitations of this approach may have contributed to the mixed evidence for a Matthew effect.

Instead of using initial test performance to select rich and poor readers, we tested for fan spread by defining rich and poor readers using exogenous child- and family-level characteristics indexing the relative magnitude of reading-related risks and inputs that a child is likely to experience during the preschool and early elementary school years. We did so for two reasons. First, the relative magnitudes of these risks and inputs are a particularly appropriate conceptualization of a child’s reading-related “wealth.” Thus, we were able to more fully account for the diverse sets of factors thought to influence the Matthew effect model’s predicted growth trajectories. Second, using these exogenous characteristics allowed us to determine whether the Matthew effect exists for those most at risk of reading disabilities.

We used growth curve modeling to test for fan spread. This allowed us to compute average reading growth trajectories for specific population subgroups defined by the exogenous background characteristics mentioned above. These epidemiological analyses revealed which groups of children became richer or poorer readers relative to typical readers over time. As such, the analyses empirically answered the question, “Matthew effects for whom?”

Despite our use of standard scores (see, e.g., Stanovich, 2000), we investigated this question in a way that was not methodologically tautological. That is, it was possible for our analyses to yield any of the following conclusions: (a) any, some, or none of the initially lowest performing groups moved further from the mean over time (measured in standard deviation units separately at the beginning and ending time periods); and (b) any, some, or none of the initially highest-performing groups moved further from the mean over time. We consistently observed significant effects when we estimated growth curve models to predict reading proficiency. Specifically, males typically began kindergarten with lower reading skills than females, and their skills grew more slowly than those of females. Family SES background was strongly and positively associated with both a child’s beginning reading skills and his/her subsequent rate of reading growth. Hispanic children entered school with significantly fewer reading skills than White children; Asians entered with significantly greater reading skills than Whites. Further, the reading skills of Black, Hispanic, and Asian children grew more slowly than those of White children. Among these, the Black children experienced the greatest increase in their gap, relative to White children.

When we combined the child- and family-level variables to define population subgroups, we observed a fan spread effect in the reading scores of five groups of low-performing children: (a) Blacks and Other Race or Ethnicity (including American Indian) males from the lowest SES quintile, (b) Black and Other Race or Ethnicity females from the lowest SES quintile, and (c) White males from the lowest SES quintile. Between the fall of kindergarten and the spring of third grade, their distance below average children’s reading level increased by 0.5-0.3 of a standard deviation. This seems a straightforward case of poor readers growing poorer.

In contrast, Asian females and males from families in the highest quintile of the SES distribution typically entered school as the highest-performing readers. In the fall of kindergarten, their composite reading scores averaged 1.16 and 1.01 standard deviations above the mean, respectively. However, these children did not grow to become increasingly more skilled than typical readers of the same age. Instead, by the spring of third grade, their average reading score had declined to about 0.6 standard deviations above the overall average. Other high-performing groups also failed to become increasingly more skilled (in standard deviation units, relative to the mean) readers. For example, White female children from the top quintile of SES families increased their average reading score by only 0.06 standard deviation, from 0.62 to 0.68 standard deviation above the mean. Other “rich” readers became less so. This was the case for Black and Other Race or Ethnicity males as well as for females from the highest SES quintile.

Thus, we found that the reading rich did not become richer. Put another way, children who entered school at relatively lower risk for having reading disabilities (e.g., high-SES Asian and White females) did not become relatively (that is, when compared to same age “typical” peers) better readers. Other studies (e.g., Leppanen et al., 2004; McCoach et al., 2006; Phillips et al., 2004) have
reported a similar finding. Yet, and unlike these other studies, we did not find that poorly skilled readers began to catch up with their peers. Instead, we found evidence for a one-sided Matthew effect. That is, relative to typical readers of the same age, children who entered school at relatively higher risk for having reading disabilities (e.g., low-SES Black males) became comparatively poorer readers over time. The difference between our findings and those of other researchers may be due to our larger sample and our use of gender, race/ethnicity, and SES to index a child’s reading growth. These factors exert strong effects (e.g., McCoach et al.). At the very least, results from our study and others’ suggest that Stanovich’s (1986) Matthew effects model may not produce uniform influences on “rich” and “poor” readers’ progress in becoming literate.

Limitations
Our study is subject to at least three limitations. First, our data were limited to students in kindergarten through third grade. Thus, we do not know the extent to which the Matthew effect continues as children move beyond the lower elementary grades. Second, we did not directly test the specific developmental model thought to cause the Matthew effect. For example, we did not test whether children at risk for reading disabilities became less motivated to engage in reading or practiced it less frequently than their peers (Chapman, Tunmer, & Prochnow, 2000; Stanovich, 1986). However, this study was not designed to empirically evaluate the Matthew effects model’s interrelated, reciprocally causative mechanisms. Instead, we investigated which, if any, population subgroups experienced the model’s predicted fan spread effect. We believe that a useful first step in evaluating the validity of Stanovich’s theoretical model is to resolve whether and for whom the predicted Matthew effect in fact occurs. Thus, and although our results indicate that a fan spread effect does occur for children from those population subgroups most at risk for later being identified reading disabled, the causal mechanisms underlying this fan spread, as well as interventions capable of reducing or eliminating it, require further study.

Third, we entered only a small set of exogenous variables into the growth curve models. We did so because previous research has indicated that differences in gender, race/ethnicity, and SES would be particularly powerful indicators of a child’s risk status (McCoach et al., 2006). This indeed proved to be the case. That is, children of different gender, race/ethnicity, and SES groups, on average, performed differently on this study’s measure of reading proficiency. However, this is not the same as saying that a child’s gender, race/ethnicity, or SES should be construed as a lasting marker of his or her status as a good or poor reader. Any given subgroup in our sample included both good and poor readers. As is always the case in these types of analyses, “risk” is probabilistic rather than deterministic.

Contributions to the Literature
Our findings contrast with those of others as to whether a Matthew effect exists (e.g., Aarnoutse & van Leeuwe, 2000; Catts et al., 2003; Jordan et al., 2002; Leppanen et al., 2004; McCoach et al., 2006; Parrila et al., 2005; Phillips et al., 2002; Shaywitz et al., 1995). For example, Shaywitz et al. used both growth modeling and statistical control for exogenous variables, but did not find evidence for fan spread between rich and poor readers.

Several factors may account for differences in our and others’ findings. First, our dataset was much larger (i.e., 10,587 children) than all others to date. For example, Scarborough and Parker’s (2003) null finding was based on a small sample of 57 children with and without learning disabilities. Thus, our sample provided ample statistical power to detect effects. Second, in our analyses, reading achievement intercepts and slopes were allowed to be a function of the child’s or family’s background variables. Third, in testing for Matthew effects for particular groups of children, we included SES. Whereas this factor may strongly influence children’s reading growth during their preschool and elementary school years, and likely captures some of the causal forces underlying the Matthew effect model on their acquisition of reading skill (e.g., Shaywitz et al.; Stanovich, 1986), it has not typically been incorporated into other investigators’ analyses (e.g., Leppanen et al.; Parrila et al.).

Implications for Practice
Our findings help identify which population subgroups are likely to lag increasingly behind in becoming proficient readers. Put another way, our findings starkly illustrate the power of a small set of background variables (i.e., gender, race, and socioeconomic class) to explain the relative reading progress – or lack of progress – of large subgroups of children in the United States. Evidence for these patterns remained after the use of rigorous statistical techniques. We conclude that a one-sided Matthew effect exists, and it exists for those socio-demographic subgroups most at risk for later being identified as having reading disabilities.

For practitioners, the implications are clear. They, as well as researchers and policy-makers also working to “leave no child behind,” should consider increasing their efforts to provide intensive and high-quality early interventions to at-risk children. Without the benefit of more intensive early intervention efforts, the groups of children who arrive in kindergarten already at risk for
being poor readers are likely to continue to fall further behind their peers. Children entering school as relatively low-skilled readers are unlikely to "catch up" with their higher-skilled peers (even as these higher-skilled peers slow their rate of skills growth) and, as such, will likely need early intervention efforts that are more intensive than those currently being employed. Certain groups of children (e.g., males being raised in low-income households) will likely be most in need of such early efforts, as these children are both at risk of beginning school as relatively less skilled readers and of falling increasingly further below typical readers as they grow older.

REFERENCES


Please address correspondence regarding this article to: Paul L. Morgan, Department of Educational Psychology, School Psychology and Special Education, 211 CEDAR Building, University Park, PA 16802; e-mail: paulmorgan@psu.edu
Copyright of Learning Disability Quarterly is the property of Council for Learning Disabilities and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.