ABSTRACT: **Purpose:** This investigation examined the performance of 50 African American children on a reading comprehension test.

**Method:** Longitudinal data were compared for two groups of students who were preschoolers or kindergartners at Time 1 and elementary-grade students at Time 2. Outcomes were examined for positive predictive relationships based on their oral language and cognitive skills as preschoolers and kindergartners at Time 1. The Time 1 preschoolers were all from low-income homes, whereas the Time 1 kindergartners were all from middle-income homes. All students were urban dwellers and speakers of African American English.

**Results:** Two measures predicted later reading comprehension levels for the Time 1 preschoolers: use of complex syntax and shape matching. The Time 1 preschoolers and kindergartners showed no significant differences in reading comprehension at the end of first grade, but the preschoolers were significantly ahead of the kindergartners in reading by third grade.

**Clinical Implications:** The potential of preschools that emphasize early language and literacy for improving the reading outcomes of African American students is discussed.

**KEY WORDS:** child language acquisition, literacy development, minority language, multicultural
early grades remain behind (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Torgesen & Burgess, 1998). It is simply too late to wait to intervene until the middle elementary years. Middle elementary school is a time when most systematic identification takes place, and then in the high-stakes context of students being unable to extract meaning from texts across the content areas. The development of effective prevention programs for African American students depends on knowing which early abilities are beneficial to later reading achievement.

A small number of strong predictors of later reading skills are emerging for the general population. Letter knowledge (Burgess & Lonigan, 1998; Hecht, Burgess, Torgesen, Wagner, & Rashotte, 2000; Scarborough, 1998) and family literacy practices, especially opportunities for home-based story book reading (Bus, 2001; Bus, van IJzendoorn, & Pellegrini, 1995), are all strong predictors for mainstream children. In particular, however, letter knowledge reflects better knowledge of and experience with print (Adams, 1990), and must be viewed with extreme caution when extrapolated to the understanding of sociocultural groups like African American children. The world experiences and home literacy practices of these children may differ in major ways from those of their mainstream peers. Further, home literacy practices may be quite different for African American families in terms of both frequency and style when compared to the practices of families of majority children (Anderson-Yockel & Haynes, 1994; Federal Interagency Forum on Child and Family Statistics, 2000; Hammer, 2001).

Phonological awareness skills are a robust predictor for reading achievement, but again, most of the research supporting this relationship has been conducted with majority children, or performances have not been disaggregated for African American participants (Bus & van IJzendoorn, 1999; Hecht et al., 2000; Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). At this time, virtually nothing is known about the phonological development of young African American students. A few studies have compared specific phonological processes shared by African American English (AAE) and Standard American English (SAE) spoken by children (Haynes & Morán, 1989; Seymour & Seymour, 1981), but in-depth, comprehensive, and systematic inquiry remains forthcoming. Thus, it would be unfortunate to repeat the mistakes of the past wherein measures developed and appropriate for SAE-speaking students are used as a standard for evaluating the linguistic skills of AAE-speaking children. In the past, incomplete information has led to test bias and consequent inappropriate and unfavorable interpretations. Examination of the impact of phonological awareness skills of AAE-speaking students on reading outcomes will no doubt be informative, but must be based on foundational characterizations of the phonological characteristics and the developmental course of child AAE.

More comprehensive information about the letter knowledge, family literacy practices, and phonological awareness skills of African American students is clearly warranted and will require an intensive and extended research effort. Are there other early predictors of reading success for African American children that are informative in their own right, and of special value now in the absence of information on letter knowledge, family literacy practices, and phonological awareness?

Language skills are particularly good candidates in the search for early predictors of reading success. Six of the 10 indicators of success or failure in reading identified by the National Research Council’s report on preventing reading difficulties in children relate to language skills (Snow, Burns, & Griffin, 1998). Positive relationships are apparent between individual oral language skills and later reading achievement, and children with impairments in oral language demonstrate more frequent and significant problems with reading acquisition (Bishop & Adams, 1990; Catts, 1993; Scarborough, 1989).

Until recently, characterizations of the language skills of African American children at the time of school entry were unavailable, precluding examination of the relationship between their early language and later reading. At the University of Michigan, data have been gathered on the early sentence structure and vocabulary skills of African American students, including average length of communication units (Craig, Washington, & Thompson-Porter, 1998a), frequency of complex syntax (Craig & Washington, 1994), and number of different words as a measure of expressive vocabulary (Craig & Washington, 2000; Craig, Washington, & Thompson-Porter, 1998b). These measures are long-standing for majority children, and when reexamined were found to be informative ways to characterize the early language skills of African American children (Craig & Washington, 2000, 2002). Features that recommended these measures for minority students were (a) all were age and/or grade sensitive, (b) their performance distributions can be quantified and characterized independently of the child’s dialect, and (c) average communication unit lengths, frequencies of complex syntax, and numbers of different words can all be sampled during low or moderately structured, child-centered language sampling contexts.

Child-centered language sampling is of particular importance when assessing minority language children to reduce the impact of assumptions derived from the majority culture on the child’s use of language.

The purpose of this investigation was to begin to search for early positive predictors of later reading achievement for African American students. Unlike phonology, foundational information on the sentence structure skills of AAE-speaking students is available (Craig & Washington, 1994, 2000, 2002; Craig et al., 1998a) and lends itself well to this inquiry. Accordingly, in order to begin to explore the relationships between early oral language skills of African American students and later reading performance, the following questions were posed using the archival database for African American students at the University of Michigan:

- Are there statistically significant relationships between early oral language scores and later reading comprehension achievement levels? Specifically, do average length of spoken units, amounts of complex syntax, average number of different spoken words, and oral responding to probes of requests for information
predict Metropolitan Achievement Tests–Seventh Edition (MAT; 1993) scores at age 9?
• Does early oral language performance predict the rate of growth in MAT scores from first through third grades?

METHOD

Design

This investigation is a retrospective longitudinal examination of the relationship between early oral language skills and later achievement on a standardized test of reading comprehension using archival data for African American students at the time of school entry. Students were included in this investigation if they had participated in an oral language assessment (Time 1), had remained in the school district, and had 1 or more years of achievement test data available. Not all schools in the district administered the standardized reading test, and those that did did not necessarily do so each year. Accordingly, 50 of the 66 preschool and kindergarten students who participated in the Time 1 testing were still enrolled in the school district at the time of this investigation and had available achievement data.

Archival achievement test data were available for more than 1 year for most of the students (47/50, 94%), yielding 127 outcome scores between first and third grades. Hierarchical linear modeling techniques (HLM; Bryk, Raudenbush, & Congdon, 1996) were applied to the data. Performances on oral language tasks at Time 1 were used to predict reading comprehension outcomes at age 9 years, as well as growth in reading comprehension scores over time.

Participants

The participants were 50 typically developing African American students living in metropolitan Detroit, Michigan. African Americans constituted more than 75% of the children enrolled in the participating school district. The children were recruited to participate in the project when they were preschoolers or kindergartners (Time 1). Recruitment was initiated by the school principals sending home a project description and consent form.

The children ranged in age at Time 1 from 4;2 (years; months) to 6;3. The gender distribution of the sample was allowed to vary (see Table 1 for the resulting distributions). Whereas all of the children were enrolled in their first year of public schooling, their grade at enrollment in public education differed by socioeconomic status (SES). In Michigan, public preschool is available only to children considered “at risk,” and a primary risk factor is low-income status. Accordingly, this state offers a natural experiment for evaluation of differences in family income and time of school entry into public education. All of the preschoolers were from low-income families and all of the kindergartners were from middle-income families. SES was confirmed from one or more of the following sources: the participants’ eligibility or ineligibility to participate in the federally funded free or reduced-price lunch program in their schools, their eligibility for a state-funded preschool program, and/or the Hollingshead Four Factor Index of Social Status (Hollingshead, 1975) derived from caregiver interviews.

All of the children were speakers of AAE but varied considerably in the degree to which AAE was used in spontaneous discourse. Table 1 presents outcomes for calculation of a dialect density measure (DDM) for each student, determined by dividing the frequencies (tokens) of AAE in picture description language samples by the number of words (tokens) in the same samples (Craig et al, 1998a; Craig & Washington, 2000, 2002). AAE tokens were identified using Washington and Craig’s criteria (1994, 2002) for a potential set of 24 morphosyntactic features produced by students of these ages. There were no statistically significant differences by grade on DDM, $t(48) = 0.54, p > .05$.

Data Collection at Time 1

Assessments were conducted in the second half of the school year for both preschoolers and kindergartners. Expressive language samples, responses to a Wh-questions task, and performances on the Triangles subtest of the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983) were collected for each student at Time 1. The tasks were randomly ordered in the data collection protocol to avoid systematic order effects in the data. The protocols were administered by African American female examiners who spoke AAE to the children. The children and examiners were unacquainted, which is consistent with most oral language assessment contexts.

Table 1. The distribution of the participant sample (N = 50) at Time 1 for grade, age (in months), and gender, and their dialect density measure (DDM) by grade.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Age</th>
<th>Gender</th>
<th>DDM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Male</td>
</tr>
<tr>
<td>Preschool</td>
<td>55.5</td>
<td>3.9</td>
<td>17</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>69.2</td>
<td>3.6</td>
<td>13</td>
</tr>
<tr>
<td>Combined</td>
<td>62.4</td>
<td>7.9</td>
<td>30</td>
</tr>
</tbody>
</table>
Expressive language samples were collected during a picture description task. Pictures were selected from the Bracken Concept Development Program (Bracken, 1986), and depicted an accident (picture #5), a winter scene (picture #7), and a school crossing guard (picture #24). These pictures were selected because they presented adults and children who were racially diverse. In addition, the individuals were involved in interesting activities for both male and female children. The examiner prompted each picture by saying, “Tell me about this picture.” If the student simply labeled objects, the examiner prompted further with “Tell me what’s happening in the picture.” Each child wore an individual microphone and the samples were audio-recorded using a microphone, mixer, and audio recorder. The pictures were presented in random order for each child, and the task was untimed.

**Analyses of Language Samples**

The language samples were transcribed orthographically using the segmentation criteria for communication units (C-units). A C-unit has been defined as independent clauses plus their modifiers in the form of coordinate, subordinate, and embedded clauses (Loban, 1976). This process required segmenting successive main clauses linked by simple coordinate conjunctions (and, but, or) into separate C-units if the second clause included a subject. When the second clause elliptically omitted the subject, the two clauses were considered a single C-unit. For example, “He fell and hurt himself” was considered a single C-unit, whereas the turn “and the car was driving/ and her put the stop sign up” was segmented into two C-units as indicated by the .

Loban’s criteria included nonclausal utterances in the C-unit corpora if they were responses to prior adult questions. Similarly, single-word stereotypical acknowledgments to prior adult comments and child productions of “what?” functioning as a contingent query were included. Other potential single word forms, particularly “wow!” or other fragments functioning as exclamations, were not included in the C-unit corpora. C-units were transcribed into Codes for the Human Analysis of Transcripts (CHAT) files consistent with the conventions of the Children’s Data Exchange System (CHILDES; MacWhinney, 1994). Only wholly intelligible C-units were included in subsequent analyses.

The samples were scored for amount of complex syntax, number of different words, and mean length of C-units. The Computerized Language Analysis (CLAN) analysis programs of CHILDES automatically generated the frequencies of complex syntax codes as well as word lists using the frequency command (FREQ). The average C-unit length in words was generated by the mean length of turn command (MLT). Each of these measures are discussed in more detail in Craig and Washington (2000) and are described briefly below.

**Syntactic complexity.** Each C-unit was scored for the presence of complex syntax (Csyn) using Craig and Washington’s (1994) taxonomy. The Appendix presents the scoring taxonomy with examples from the picture description analyses. Types of complex syntax ranged from simple infinitival relationships (e.g., “she tryin(g) to get him out” coded as an infinitive with same subject) to more complex relative clauses modifying a noun or pronoun (e.g., “I see a boy who dropped the paper”), and clauses joined by a variety of more cognitively advanced subordinate conjunctions (e.g., “that’s when they got hurt (be-cause they was ride-in(g) they bike”). More than one complex syntax code was possible per C-unit, as in the prior example. Only connectives internal to the C-unit were scored as conjunctions for the purposes of this complex syntax analysis. Connectives that were in the initial word position of a C-unit were not considered an example of complex syntax. An example of a turn internal connective that linked consecutive C-units but was not scored for complex syntax was “and a kid he is in the street/ and a kid dropped his homework.” A ratio depicting the amount of complex syntax was calculated by dividing the tokens of complex syntax by the number of C-units in the sample. This rate was then multiplied by 100 to provide a range of scores sufficient for statistical treatments.

**Expressive vocabulary.** As a measure of lexical diversity and expressive vocabulary, the number of different words (NDW) used in the picture description samples was calculated (Craig & Washington, 2000; Watkins, Kelly, Harbers, & Hollis, 1995). The FREQ command of CLAN automatically generated word lists. These lists were then edited so that morphological variations in the form of number and tense markers on regular nouns and verbs were ignored. Irregular forms of nouns and verbs, for example, leaf, leaves, and do, did, were treated as separate lexical forms. However, regular forms of nouns and verbs, for example, boy, boys, and wants, wanted, and wanting, were treated as the same noun or verb lexical root and were not scored as different words although the FREQ command displays them as different types. Although these bound morphemes may be variably included or excluded in AAE, it seemed appropriate to represent this aspect of production within the dialect computations rather than the analysis of different words.

**Average C-unit lengths.** The mean length of the C-units (MLCU) in words was calculated for each participant’s picture description samples.

**Comprehension of requests for information.** This task presented two activity pictures depicting barbecuing and snow shoveling and probed student’s responses to questions. Each picture was accompanied by 12 questions presented by the examiner using AAE. These questions included: What this (is)? What he doin’? Who (object) this? Who this? How many (objects) in the picture? Where this? How long will it take to (perform action)? Why he (perform action)? How far he (perform action)? How he (perform action)? How often he (perform action)? and, When this happenin’? Although pilot work indicated that order effects were not apparent for nonrandomized trials

---

1 The stimuli were two pictures developed by the authors or two selected from the Bracken Concept Development Program (cards 33 and 35; Bracken, 1986). Earlier work (Craig et al., 1998b) found no differences in performance between the two sources of materials and either are recommended to the reader for these purposes. Copies of the authors’ pictures and the computerized scoring software are available from the authors.
Assessment of Achievement

A computerized scoring program generated total scores and total possible score was 72 points (see Craig et al., 1998b). The scoring system used assigned full credit (3 points) if the child produces the target response, 2 points if the child responds to the pragmatic intent of the specific request for information but uses a nonspecific referent or misnames the referent (e.g., Examiner: “How he movin’ the snow?” Student: “With a lawnmower;”), 1 point if the child responds but to a potentially different question (e.g., Examiner: “How often they barbecue?” Student: “Because they hungry”), and 0 points if the child says something unrelated, says “I don’t know,” or does not respond. The total possible score was 72 points (see Craig et al., 1998b).

Assessment of Nonverbal Cognition

In addition, each child was administered the Triangles subtest of the K-ABC (Kauffman & Kaufman, 1983). The Triangles subtest is a matching task that reportedly taps a fairly general nonverbal cognitive skill, is appropriate for children in this age range, and evidences no racial or cultural biases (Cole, Gay, Glick, & Sharp, 1971; Kaufman, 1973; Lempley & Rust, 1986; Palmer, Olivarez, Willson, & Fordyce, 1989; Willson, Nolan, Reynolds, & Kamphaus, 1989). Equivalencies on the cognitive task at Time 1 were examined by comparing the scaled scores on the Triangles subtest. The preschoolers and kindergartners at Time 1 were not significantly different on their Triangles subtest performances, t(48) = –4.10, p < .001. The scaled scores have the advantage of controlling for age-related effects between the preschoolers and kindergartners, and thus allowed a better examination of cognitive status.

Assessment of Achievement

Scaled scores on the reading comprehension portion of the MAT (1993) were provided for each student by the school district, when schools administered the test, it was meaningful to support learning across the academic content areas. Age was used rather than grade because the MAT scaled scores are related to age. Variables that were categorical and bivariate, specifically gender, SES, and grade, were entered into the models as dummy coded variables. For example, for SES, children from low-SES (LSES) homes were coded as 0 and those from middle-SES (MSES) homes were coded as 1 for analysis purposes. Gender did not significantly contribute to either intercept, T-ratio(117) = 1.10, p > .05, or slope, T-ratio(117) = –.50, p > .05, and was removed from the linear equation.

Scoring Reliability

Reliabilities were established for the picture description samples by randomly selecting one of the three pictures in each transcript and having an independent observer re-transcribe it. Hence, approximately one third of the data were analyzed for reliability purposes. For transcription, a point-to-point comparison at the level of the morpheme was high (90%) when the number of morphemes in agreement was divided by the number of agreements plus disagreements. C-unit segmentation for these same samples was also high (93%).

Five transcripts, representing approximately 10% of the corpus, were randomly selected and the complex syntax recoded by an independent observer. Point-to-point comparisons were calculated by dividing the number of agreements by the number of agreements plus disagreements. The percentage of agreement was 87% for tokens and 100% for types, indicating high levels of agreement across transcription and scoring portions of the data reduction.

Statistical Analyses

Scores at Time 1, later achievement outcomes, and their interrelationships were examined using HLM (Bryk et al., 1996). Multilevel analysis with HLM allowed for the examination of increases in skill over time (growth curves) while controlling for individual child characteristics that had the potential to influence the outcomes. HLM is especially well-suited for the analysis of this type of data because statistically, it manages nested variables as well as differences in time of entry into the longitudinal design (preschool vs. kindergarten).

Using MAT as outcome data, a two-level HLM was fitted to the age distributions. Specifically, the HLM process was threefold: (1) building the best unconditional model; (2) examining the effects of cognition and each of the four oral language measures on both the intercept and the slope of the MAT outcomes by running each separately to avoid a Type I error given the relatively small sample sizes; and then (3) running a final version of the model with the Time 1 variables that had been found to contribute, at statistically significant levels, to the MAT outcomes. To do this, MAT outcomes were anchored at age 9 years. Although the endpoints for growth curves are somewhat arbitrary, age 9 was selected because it corresponded to the third grade, the time when reading skill must be adequate. Variables that were categorical and bivariate, specifically gender, SES, and grade, were entered into the models as dummy coded variables. For example, for SES, children from low-SES (LSES) homes were coded as 0 and those from middle-SES (MSES) homes were coded as 1 for analysis purposes. Gender did not significantly contribute to either intercept, T-ratio(117) = 1.10, p > .05, or slope, T-ratio(117) = –.50, p > .05, and was removed from the linear equation.

RESULTS

The data were examined in three ways. First, the similarities and differences between the preschoolers and the kindergartners at the time of the initial testing were examined. Second, the performances of the two groups on the reading comprehension portion of the MAT were
examined across grades one, two, and three. Finally, the contributions of the early measures to explaining the variance in the later MAT achievement data were examined, both in terms of the levels of reading comprehension achieved and in terms of the rate of growth in obtaining these levels. Each aspect of the data is discussed below.

Early Oral Language Status

The preschoolers and kindergartners demonstrated similar performances on each of the oral language measures with the exception of the responses to Wh-questions task. The Wh-questions task elicited significantly higher point scores for the kindergartners (M = 59) as compared to the preschoolers (M = 53), independent t(48) = –4.10, p < .001 (see Table 2). Whereas the kindergartners were chronologically older than the preschoolers, not surprisingly, both significant and nonsignificant differences between the preschoolers and the kindergartners on the oral language measures favored the kindergartners.

Later Reading Comprehension Performance

Figure 1 displays the growth in MAT reading comprehension scores as a function of age. Not surprisingly, levels in reading comprehension as measured by the MAT scaled scores were significant over time for both preschoolers, T-ratio(48) = 49.78, p < .001, and kindergartners, T-ratio(48) = 56.77, p < .001. Growth in MAT reading comprehension scores was statistically significant for preschoolers, T-ratio(48) = 6.16, p < .001, and for kindergartners, T-ratio(48) = 7.61, p < .001 as well. On the one hand, these findings simply indicate that reading comprehension increased steadily as the students progressed through the elementary grades. However, the linear trends were significant for preschoolers, t(48) = 6.16, p < .001 and for kindergartners, T-ratio(48) = 3.80, p < .001, indicating that scores improved over time at a consistent rate. The quadratic trends were nonsignificant for both groups—T-ratio(121) = .28, p > .05, for preschoolers, T-ratio(121) = –1.6, p = .12, for kindergartners—providing further support for the linear relationship; in other words, change in scores over time was neither accelerating or decelerating. Further, this growth in reading comprehension was significantly greater for the preschoolers as compared to the kindergartners, slope difference = 6.68, T-ratio(44) = 5.10, p < .001, see Table 3. In the absence of quadratic trends, the findings as a whole indicate that growth in reading comprehension had not tapered off by age 9 and that this growth was faster for the preschoolers.

The mean MAT reading comprehension scaled scores are presented in Table 4 for each grade, along with the percentile ranks and the grade equivalent scores that correspond to each mean. Whereas MAT testing was conducted in May of each academic year, the grade equivalent scores should approximate 1;8 (years;months), 2;8, and 3;8 for each grade respectively. The data indicate that after first grade, the kindergartners began to fall behind grade expectations. The first graders on average obtained grade equivalents comparable to their time of testing (i.e., May of their first grade). At second and third grades, however, the kindergartners evidenced a cumulative lag, both behind the preschooler cohort and from grade expectations. At first grade, the fitted MAT scaled scores were not significantly different between Time 1 preschoolers and kindergartners, t(48) = 1.76, p > .05.

Table 2. Means (M) and standard deviations (SD) for percentages of complex syntax (Csyn), numbers of different words (NDW), mean length of C-units (MLCU), responses to Wh-questions (Wh-q), and the Triangles scaled score (TRI) at Time 1 by grade (N = 50).

<table>
<thead>
<tr>
<th>Grade</th>
<th>Csyn</th>
<th>NDW</th>
<th>MLCU</th>
<th>Wh-q</th>
<th>TRI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Preschool</td>
<td>18</td>
<td>11</td>
<td>69</td>
<td>18</td>
<td>4.24</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>23</td>
<td>19</td>
<td>70</td>
<td>24</td>
<td>4.80</td>
</tr>
<tr>
<td>Combined</td>
<td>21</td>
<td>15</td>
<td>69</td>
<td>21</td>
<td>4.52</td>
</tr>
</tbody>
</table>

Note. na = not appropriate.
However, at both second, \( t(48) = 2.79, p < .001 \), and third, \( t(48) = 3.67, p < .001 \), grades, the cohorts performed significantly differently from each other, with the advantage to the preschoolers.

### The Effects of the Oral Language and Nonverbal Cognition Measures on MAT Outcomes

Three of the oral language measures—NDW, MLCU, and Wh-q—showed nonsignificant relationships to the MAT outcome measures. NDW did not significantly affect children’s MAT reading comprehension intercepts, preschoolers: \( T_{ratio}(46) = -0.83, p > .05 \), kindergartners: \( T_{ratio}(46) = 0.45, p > .05 \), or slopes, preschoolers: \( T_{ratio}(46) = 0.86, p > .05 \), kindergartners: \( T_{ratio}(46) = 0.35, p > .05 \). MLCU did not significantly affect children’s MAT reading comprehension intercepts, preschoolers: \( T_{ratio}(46) = 1.11, p > .05 \), kindergartners: \( T_{ratio}(46) = -1.06, p > .05 \), or slopes, preschoolers: \( T_{ratio}(46) = 0.06, p > .05 \), kindergartners: \( T_{ratio}(46) = -0.70, p > .05 \). Like

Table 3. Hierarchical linear modeling results for Metropolitan Achievement Tests reading comprehension scores examining the effect of Triangles (TRI) and percentage rate of complex syntax use (Csyn) centered at the grade group mean by grade at Time 1.

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>T-ratio</th>
<th>( p ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi_g )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preschool</td>
<td>( \beta_{g0} )</td>
<td>595.57</td>
<td>10.62</td>
<td>56.09</td>
</tr>
<tr>
<td>TRI</td>
<td>( \beta_{g3} )</td>
<td>13.35</td>
<td>3.30</td>
<td>4.05</td>
</tr>
<tr>
<td>Csyn</td>
<td>( \beta_{g4} )</td>
<td>2.40</td>
<td>1.03</td>
<td>2.32</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>( \beta_{g0} )</td>
<td>570.57</td>
<td>8.26</td>
<td>69.08</td>
</tr>
<tr>
<td>TRI</td>
<td>( \beta_{g3} )</td>
<td>2.16</td>
<td>2.67</td>
<td>0.81</td>
</tr>
<tr>
<td>Csyn</td>
<td>( \beta_{g4} )</td>
<td>-0.25</td>
<td>0.53</td>
<td>-0.49</td>
</tr>
</tbody>
</table>

| For slope    |             |                |         |              |
| \( \pi_1 \)  |             |                |         |              |
| Preschool    | \( \beta_{11} \) | 34.95          | 4.92    | 7.10         | .000 |
| TRI          | \( \beta_{13} \) | 4.15           | 1.57    | 2.64         | .012 |
| Csyn         | \( \beta_{14} \) | 0.57           | 0.47    | 1.22         | .229 |
| Kindergarten | \( \beta_{12} \) | 28.27          | 3.44    | 8.20         | .000 |
| TRI          | \( \beta_{15} \) | -0.43          | 1.10    | -0.39        | .695 |
| Csyn         | \( \beta_{16} \) | -0.11          | 0.24    | -0.46        | .647 |

#### Note.
Deviance = 1236.45; \( df \) at level 1 = 44. The \( \chi^2 \) statistics reported are based on only 47 of 50 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Table 4. Mean Metropolitan Achievement Tests reading comprehension outcomes in scaled scores by preschool (Pre) and kindergarten (K) cohort at Time 1 and the percentile ranks and grade equivalents corresponding to each mean.

<table>
<thead>
<tr>
<th>Grade equivalent</th>
<th>Grade</th>
<th>M</th>
<th>SD*</th>
<th>Minimum</th>
<th>Maximum</th>
<th>N</th>
<th>Percentile rank</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pre</td>
<td>526*</td>
<td>34.1*</td>
<td>454</td>
<td>585</td>
<td>25</td>
<td>63</td>
<td>2;1</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>514*</td>
<td>34.1*</td>
<td>447</td>
<td>578</td>
<td>25</td>
<td>53</td>
<td>1;9</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pre</td>
<td>561**</td>
<td>34.1*</td>
<td>483</td>
<td>638</td>
<td>23</td>
<td>43</td>
<td>2;7</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>542**</td>
<td>34.1*</td>
<td>478</td>
<td>654</td>
<td>24</td>
<td>29</td>
<td>2;4</td>
</tr>
<tr>
<td></td>
<td>Third</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pre</td>
<td>596**</td>
<td>34.1*</td>
<td>529</td>
<td>731</td>
<td>9</td>
<td>54</td>
<td>4;1</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>571**</td>
<td>34.1*</td>
<td>499</td>
<td>624</td>
<td>21</td>
<td>32</td>
<td>2;9</td>
</tr>
</tbody>
</table>

* Calculated using the final hierarchical linear modeling results.
* Significant at \( p < .05 \).
NDW and MLCU, Wh-q did not significantly affect children’s MAT reading comprehension intercepts, preschoolers: T-ratio(46) = 1.91, p = .06, kindergartners: T-ratio(46) = 1.35, p = .18, or slopes, preschoolers: T-ratio(46) = 1.43, p = .16, kindergartners: T-ratio(46) = 0.40, p = .69.

In contrast to the other measures, Csyn and Triangles evidenced significant relationships to MAT outcomes.\(^2\) Accordingly, the model was run again, this time with both Csyn and Triangles included in the model (see Table 3). NDW, MLCU, and Wh-q were not included in this final model. Csyn evidenced a significant effect on the intercept for the MAT scores at age 9 and the effect was large, T-ratio(44) = 2.33, p = .025, effect size for 1 SD change = 1.05 for preschoolers at Time 1, but not for kindergartners at Time 1, T-ratio(44) = –0.49, p > .05, for kindergartners, T-ratio(44) = –0.46, p > .05, at Time 1, indicating that although achievement level was impacted systematically by preschooler syntactic skill, growth in reading comprehension was not (see Table 3).

Performance on the Triangles subtest significantly contributed to the MAT comprehension for preschoolers’ intercepts, and this effect was large, T-ratio(44) = 4.05, p < .001, effect size for 1 SD change = .98, and slopes, T-ratio(44) = 2.65, p = .012. This coefficient indicated that for every one point increase in the raw score on the Triangles subtest, the MAT scaled scores increased by 13.4 scaled score points (see Table 3, \(\beta_{\text{Tri}}\)). Triangles did not contribute significantly to the intercepts, T-ratio(44) = 0.81, p > .05, or slopes, T-ratio(44) = –0.39, p > .05, of the kindergartners. In other words, the Triangles subtest scores were related to reading comprehension outcomes for preschoolers, predicting both achievement level and learning progress, but were not predictive for kindergartners.

### DISCUSSION

This investigation examined the relationship between the oral language skills of preschool and kindergarten African American students and their later performance on a standardized test of reading comprehension. The two major findings were as follows.

1. Statistical modeling revealed that performance on the MAT increased significantly across the time period studied, from first through third grade. There were important differences between the achievement levels of the Time 1 preschoolers and kindergartners.

2. Performance on the Triangles subtest and amounts of complex syntax positively predicted MAT outcomes for Time 1 preschoolers.

### Finding 1: Increases in MAT Performance

Rate of growth in scores increased significantly faster for students in the Time 1 preschool cohort as compared to the kindergartners. Grade level expectations were achieved and maintained by the Time 1 preschoolers between first and third grades. In contrast, by third grade, performance for the Time 1 kindergartners was one grade level behind. Although the reading comprehension scores of the Time 1 kindergartners were commensurate with expectations at first grade, performance steadily slipped across the testing administrations in the next 2 years, relative to the nationally standardized comparison sample undergirding MAT interpretations. These performance lags are consistent with the national trends for African American students and demonstrate the impact of the Black–White achievement gap at the local level. Unlike the national picture, the students in this investigation were not behind performance expectations at the end of first grade. Their average scores on the MAT corresponded to grade equivalents that matched their time of testing (1:8). This finding was unanticipated. Nationally, African American students begin formal education behind expectations, and remain behind (Donahue et al., 2001). What factors mitigated the achievement gap for the current students at first grade?

The current students consisted of two cohorts who had many similarities and differences. In particular, they all lived in the same large urban center and attended the same schools. One noticeable difference was that all of the Time 1 preschoolers were from LSES homes and had been enrolled in a state-funded public preschool program for children at risk for academic failure. In contrast, the Time 1 kindergartners were all from MSES homes, and by implication entered school with the advantages that residence in a family and community with significant resources offers any student. It is not surprising, therefore, that the kindergartners were performing at expected levels early in the course of their formal schooling, at first grade. The nature of the early experiences of the MSES kindergartners before their enrollment in public education was unknown to the project, but likely these experiences varied across children. Some or all of the kindergartners may have had day care or private preschool experience prior to enrollment in kindergarten. If they did, the differences obtained between groups is even more striking. Whatever experiences the MSES kindergartners did have, they were not as effective as those provided to the preschoolers by the public programs.

The performance of the preschoolers is what is striking and needs explanation. The state of Michigan where the investigation was conducted provides public preschool experience for children considered at risk for academic failure through the Michigan School Readiness Program.
(MSRP), and these preschoolers were enrolled in this program. A major focus of the MSRP is an emphasis on early literacy development. Accordingly, the LSES preschoolers and the MSES kindergartners at Time 1 offered a natural experiment on the positive impact that public preschool emphasizing early literacy skills can have on later achievement.

At first grade, the children in the two cohorts were not significantly different statistically on the MAT, regardless of the SES status of their families. The present data suggest, therefore, that the public preschool experiences were effective in minimizing the differences in literacy between African American children from LSES and MSES homes. By implication, the findings that African American students enter schools already behind their majority peers (Francis et al., 1996; Torgesen & Burgess, 1998) may reflect that most are not from MSES homes, and if these children are from LSES homes, most may not have had the advantage of a formal preschool preparation as represented by public schooling at the preschool level. Considered together, the findings are a strong positive indicator that participation of preschoolers in public school programs for children at risk can equalize the starting points for all African American children at first grade, and that the negative effects on literacy of poverty and its covariables can be overcome. The present findings are suggestive that current federal initiatives that attempt to infuse language and literacy into preschools, in particular, Early Reading First (No Child Left Behind Act, 2001), are well-advised.

Of import, although the LSES children with preschool experience were functioning in similar ways on the MAT compared to the MSES kindergartners when assessed at first grade, the two groups did not perform similarly at 9 years. The children with public preschool experience performed significantly better. Indeed, the average MAT score for the Time 1 preschoolers at age 9 was 596, corresponding to the 54th percentile, whereas the average MAT score for the Time 1 kindergartners was 571, corresponding to the 32nd percentile. In other words, not only were the MAT scores significantly better for the Time 1 preschoolers at 9 years, but they were comparable to the standardization sample. Accordingly, this cohort of students was not a part of the Black–White test score gap.

Reynolds, Temple, Robertson, and Mann (2001) observed that model demonstration programs for preschoolers are not unique in producing positive long-term effects for low-income minority students. State and federal programs in public schools increase high school graduation rates and lower levels of juvenile crime. The present findings indicate that some of these positive impacts on high school completion and lower levels of crime may be rooted in the positive effects of very early public preschool experiences. The present data are suggestive that public preschool experiences for at-risk African American children can have measurable effects on their reading development both as they enter the elementary grades and relative to long-term outcomes. Considered together, the benefits are immediate as well as durable.

For the Time 1 kindergartners, the loss of the equivalent of approximately one grade level on the MAT between second and third grades is staggering. It is not clear whether the steep decline across these grades is anomalous or whether this trend continues unabated in the later years. It may be that the second to third grade transition was particularly challenging because of the switch in reading emphasis from decoding to comprehension that took place in the participating school districts in this investigation during these grades, as in most public schools across the nation (Chall, 1967). These interpretations are not mutually exclusive, however, and regardless of the nature of new demands for reading, on average, the Time 1 kindergartners lost ground. This interpretation is supported by the national statistics evidencing early and persistent achievement gaps from the time of school entry through 12th grade (Donahue et al., 2001; Federal Interagency Forum on Child and Family Statistics, 2000).

The low MAT scores at third grade highlight the need to attend to both growth in reading skills and absolute achievement levels. In the present investigation, these factors were represented by the slope across MAT administrations and the intercept at 9 years, respectively. It is the case that both student cohorts showed significant growth in reading across grades, but the absolute achievement levels and growth of the Time 1 kindergartners lagged behind. For African American students to achieve, and to narrow the achievement gap, prevention programs must find ways to help these students meet the performance levels necessary to access the grade level demands of the curriculum. Torgesen (2000) made this point previously. He argued that not just growth but performance within normal limits must be the measures of reading achievement. He proposed that performance above the 30th percentile should be used as a minimum standard in the absence of absolute performance standards when evaluating the effectiveness of early prevention programs for reading. The current findings provide strong support for Torgesen’s view. Improved performance in reading is incomplete unless it is considered in the context of the ultimate levels achieved.

Finding 2: Predictors of MAT Outcomes

Two measures at the time of school entry—amounts of complex syntax in spontaneous discourse and Triangles scores on the K-ABC—predicted later reading comprehension skills on the MAT. First, however, these relationships demonstrate that not all early measures of oral language predicted later reading performances. It is beyond the scope of this investigation to determine why significant relationships between MLCU, NDW, and responses to Wh-questions and later reading comprehension scores were not apparent. Perhaps the absence of significant relationships reflected a general disassociation for these measures, or perhaps effects are grade sensitive and not pertinent to the grades evaluated in this investigation (Bronfenbrenner, 1986). For instance, Vellutino, Scanlon, and Tantzman (1994) argued that phonology, semantics, and syntax are all likely to be of importance for reading development but carry different weights at different stages. The finding that measures of syntax and visual matching skills on a task...
like Triangles are important predictors for preschoolers but
not kindergartners may be another example of these
differential impacts by grade. Strong linguistic skills, like
syntax, at certain grades may buffer typically developing
African American students from failures in reading compre-
hension. Alternatively, however, the significant relationships
that were observed for complex syntax and Triangles may
mean that these skills are more critical for children residing
in LSES homes as compared to those in MSES homes.
Additional research will be needed to evaluate the merits of
these various interpretations.

Rego and Bryant (1993) and Torgesen et al. (1999)
observed that skills predicting word identification and those
predicting reading comprehension may not be the same.
General verbal ability predicted later comprehension in their
studies. The present findings suggest that syntactic maturity
may be a key component skill for measures of general verbal
ability associated with positive outcomes for reading
comprehension. Assigning an important predictive role to
syntactic skill has considerable intuitive appeal because
syntactic structure provides an interpretive discourse context
for children as they attempt to identify unfamiliar words
(Gleitman & Gillette, 1999), and because texts employ a
diverse and dense set of syntactic structures compared to
spoken discourse (Bus, 2001; Vellutino & Scanlon, 2001).

In addition to complex syntax, the other significant early
predictor of later MAT reading comprehension for the Time
1 preschoolers in this study was performance on the
nonlinguistic Triangles subtest of the K-ABC. In practical
terms, for each percent of change in complex syntax, MAT
scaled scores increased by 2.5 points, and for each raw
score change on Triangles, MAT scaled scores increased by
13.4. Consider the case of one of the boys who participated
in this investigation. At 9 years, he achieved a scaled score
of 505, corresponding to a percentile rank of 2. The
findings of this investigation indicate that improving this
child’s performance up to the 30th percentile (as recom-
manded by Torgesen, 2000) would require increasing his
MAT score by 54 points. This might be accomplished by
increasing his early syntax production level or nonlinguistic
matching skills, requiring an increase of 23 percentage
points, or a 4-point increase on Csyn or Triangles, respec-
tively. The process of developing appropriate and effective
eyear preventing programs for African American children
might benefit from targeting oral complex syntax skills and
abstract pattern matching. This is not to imply that there is
a direct and exclusive causal relationship between these
two predictive measures and later reading comprehension,
simply that as researchers and practitioners begin to search
for skill sets that lay a foundation for literacy acquisition,
complex syntax and abstract pattern matching should be
considered for their potential roles.

The Triangles subtest was selected for use in this
investigation because it is one of the few measures of
general nonverbal cognition that is culture fair for African
American children. Triangles was included to ensure that
the two cohorts of students did not differ from each other
in major ways in terms of general cognitive skill, and the
group scores at Time 1 indicated that they did not. Some
possible explanations for this relationship between

**CONCLUSION**

The findings of this investigation are only a very
preliminary step in improving our knowledge of the skills
that African American children bring to schooling and the potential relationships of these skills to later achievement. When a fuller complement of specific early linguistic skills that are positive predictors have been identified, it should be possible to use this information to develop culturally appropriate and pertinent programs of prevention for African American children.

ACKNOWLEDGMENTS

This investigation was supported by the Center for Improvement of Early Reading Achievement (CIERA) at the University of Michigan—U.S. Department of Education, Office of Educational Research and Improvement, Grant R305R70004; U.S. Department of Education, Office of Educational Research and Improvement Research, Grant R305T990368; and Research Grant 1 R01 DC 02313-01A1 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. The authors thank the participating administrators, teachers, parents, and students, and Stephen Raudenbush for helpful comments on the statistical aspects of the study.

REFERENCES


Hollingshead, A. B. (1975). Four Factor Index of Social Status. Unpublished manuscript, Yale University, New Haven, CT.


APPENDIX. THE COMPLEX SYNTAX CODING TAXONOMY

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples from the transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple infinitive with same subject</td>
<td>“kids trying to help each other up”</td>
</tr>
<tr>
<td>Simple non-infinitive wh-clause</td>
<td>“and kids telling they what they did for Christmas”</td>
</tr>
<tr>
<td>Noun phrase complement</td>
<td>“he says stop”</td>
</tr>
<tr>
<td>Let(s)/lemme and infinitive</td>
<td>“let’s see”</td>
</tr>
<tr>
<td>Relative clause</td>
<td>“a man who fell on his bike”</td>
</tr>
<tr>
<td>Infinitive with a different subject</td>
<td>“and some people are making a fire for the snow to melt”</td>
</tr>
<tr>
<td>Unmarked infinitive</td>
<td>“he make these fall”</td>
</tr>
<tr>
<td>Wh-infinitive clause</td>
<td>“they don’t know how to skate”</td>
</tr>
<tr>
<td>Gerunds and participles</td>
<td>“he tired of working”</td>
</tr>
<tr>
<td>Clauses joined by coordinate conjunctions</td>
<td>“fell off him bike and broke it”</td>
</tr>
<tr>
<td>and subordinate conjunctions</td>
<td>“the people are getting in the hospital because they hurted themself”</td>
</tr>
</tbody>
</table>

**Note.** Tag questions (“She got new clothes, don’t she?”) were part of the Craig and Washington (1994) taxonomy but did not occur in the current picture description samples.