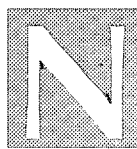


Performance of Elementary-Grade African American Students on the Gray Oral Reading Tests

Holly K. Craig
 Connie A. Thompson
 Julie A. Washington
 Stephanie L. Potter

University of Michigan, Ann Arbor



nationally, the reading skills of many African American students fail to achieve basic levels. The National Assessment of Educational Progress (NAEP; Donahue, Finnegan, Lutkus, Allen, &

ABSTRACT: Purpose: African American students perform disproportionately more poorly on standardized reading assessments than their majority peers. Poor reading performances may be related to test biases inherent in standardized reading instruments. The purpose of this investigation was to examine the appropriateness of the Gray Oral Reading Tests—Third Edition (GORT-3; Wiederholt & Bryant, 1992) for assessing the reading abilities of elementary-grade African American students.

Method: Performances of 65 typically developing African American second through fifth graders were examined on the GORT-3.

Results: African American English (AAE) was produced by most students while reading passages from the GORT-3 that were written in Standard American English (SAE). A scoring correction for AAE resulted in a statistical improvement in the performance distributions, but this did not appear to be educationally significant. Measures of total feature production predicted reading accuracy and rate, but not comprehension.

Clinical Implications: Findings are discussed in terms of the appropriateness of this instrument for use by speech-language pathologists as they contribute to curricular and classroom placement decisions in schools with large numbers of typically developing African American students.

KEY WORDS: African American students, oral reading tests, elementary-grade students, reading fluency, reading comprehension

Campbell, 2001) reported that 63% of African American fourth graders read below basic levels as compared to 27% of majority students. And, failure to attain basic levels of reading competence contributes to lower levels of academic achievement.

Several factors have been identified as contributors to the low literacy levels of African American students (for recent reviews, see Craig & Washington, in press; Washington, 2001). It is likely that no single factor in isolation is responsible for the difficulties that are faced by many African American students learning to read. Poverty and risk factors associated with low socioeconomic status (SES) play a contributing role. African American children are three times more likely than mainstream peers to live in poverty (Federal Interagency Forum on Child and Family Statistics, 2000), and poverty and low levels of literacy co-occur (Fishback & Baskin, 1991). Poverty, however, is not a sufficient explanation for reading failure among African American children because some children from low-income homes can demonstrate reading skills at better than average levels (Chall, Jacobs, & Baldwin, 1990). Moreover, poor reading performances characterize some African American children from middle-SES (MSES) backgrounds as well. Singham (1998) found that African American children from middle-SES homes scored significantly lower on a standardized measure of achievement than their majority peers who lived in the same community. Given Singham's findings, additional variables, such as the dialect spoken by African American children, need to be explored when attempting to determine the factors that contribute to low literacy levels of African American children.

Many elementary-grade African American students speak African American English (AAE; Craig & Washington,

2002). AAE is a systematic, rule-governed linguistic variety of English (Stockman, 1996; Washington & Craig, 1994; Wolfram & Fasold, 1974). AAE morphosyntactic, phonological, lexical, prosodic, and discourse features differ considerably from those of Standard American English (SAE), which is the language of the classroom and curriculum (Dillard, 1972; Hester, 1996; Labov, 1998; Morgan, 1988; Smitherman, 1998; Wolfram, 1994). This incongruity between the home language and the language used within the academic context has the potential to negatively affect the ability of African American students to become competent readers (Delpit, 1995; Ladson-Billings, 1994). Teachers, who may not be familiar with the morphological and phonological rules associated with AAE, may treat dialectal variations as isolated reading errors (Cunningham, 1976-77; Markham, 1984). For example, a common AAE phonological feature is substitution of /d/ for /ð/, as in the production of /de/ for "they." Ideally, within a teaching context, a student's production of an AAE feature during oral reading should trigger the teacher's awareness that the child is an AAE speaker who requires increased awareness of the different oral and written English dialects that are used in his or her community and school. Treating dialectal variations as reading errors may interfere with both reading instruction and the child's ability to learn correspondences between spoken phonemes and written graphemes.

Much of the early research on the relationship between reading and AAE has examined relationships between a small number of isolated AAE morphological and phonological features and performances on reading measures. These findings failed to reveal a systematic relationship between AAE productions and reading outcomes (Gemake, 1981; Goodman & Buck, 1973; Harber, 1977; Hart, Guthrie, & Winfield, 1980; Melmed, 1973; Steffensen, Reynolds, McClure, & Guthrie, 1982). In the context of so much agreement and primarily nonsignificant findings, the search for an AAE feature/reading linkage was largely abandoned. However, as the result of the persistence of disproportionately poor reading performances of African Americans, researchers have demonstrated renewed interest in potential relationships between AAE and reading achievement (Ball, 1994; Craig & Washington, 2003; Labov, Baker, Bullock, Ross, & Brown, 1998). Recently, Craig and Washington (2003) examined the AAE productions of preschoolers, kindergartners, and elementary-grade students. They observed a significant downward shift in dialect production rates between kindergarten and first grade. This shift was related to better scores on tests of reading achievement and oral vocabulary. This finding suggests that dialect density may be more useful than isolated feature use in understanding an AAE feature/reading linkage.

Fundamental to improving our understanding of AAE/reading linkages will be the ability to assess reading skills of African American students using valid measures. In order to measure the reading abilities of African American students and to plan effective instruction, it will be important to evaluate the appropriateness of the instruments that are currently used to assess reading. Prior research has shown that many cognitive and language tests developed

for majority students lack validity and exhibit bias toward African Americans (Arnold & Reed, 1976; Taylor & Payne, 1983; Vaughn-Cooke, 1986; Washington & Craig, 1992; Wiener, Lewnau, & Erway, 1983). Unfortunately, the extant reading literature includes few research studies that have investigated the potential bias of standardized reading tests against African American students.

Harber (1982) investigated the effect of crediting AAE productions of African American students' scores on the Gray Oral Reading Test (GORT, Form C; Gray & Robinson, 1967). She scored students' performances on the GORT two ways. First, all variations from print were considered errors based on directions in the manual. Second, those variations from print that were consistent with AAE features available at the time for adult speakers of AAE were not scored as errors. She found that the second set of scores significantly improved with the crediting procedure. Most variations were attributable to the use of AAE and involved the phonological feature system. In comparison to SAE-speaking peers, the performances of the African American students were significantly lower even with the scoring adjustment. In addition, the African American students read more slowly than the mainstream children. Harber recommended that reading tests be modified so that African American students would not be penalized for the use of dialect. She also recommended that future research evaluate the validity of the GORT for assessing the reading skills of African American students.

With persistently low academic achievement by many African American students, it seems important to revisit the issues Harber (1982) raised. If an updated version of the GORT (with or without a scoring adjustment) is found to be a culturally fair assessment instrument for African American students, it will be beneficial for assisting speech-language pathologists (SLPs) as they meet increasing expectations for their roles in reading acquisition. The roles of school-based SLPs have expanded to include not only the identification, assessment, and intervention of students at risk for communication disorders, but also the contribution of instructional planning, curriculum development, advocacy, professional preparation, and research in the general education context (American Speech-Language-Hearing Association, 2001). In order to accomplish these goals, research needs to study the performances of African American students on standardized reading assessments. Whereas the GORT has been revised since the early 1980s, it will be important to determine whether a scoring adjustment for AAE is still necessary and whether it is an appropriate assessment of the oral reading ability of African American students. It will be informative to compare performances of African American students on this test to their performances on other reading tests as well.

The purpose of this investigation, therefore, was to examine the appropriateness of the Gray Oral Reading Tests—Third Edition (GORT-3; Wiederholt & Bryant, 1992) for assessing the reading achievement of elementary-grade African American students. The following questions were posed:

- Do African American students produce measurable amounts of AAE during oral reading?

- Are there systematic effects of SES, grade, and gender on dialect density measures (DDMs) and reading outcomes?
- What are the performance distributions of typically developing African American students on the GORT-3?
- How do performances on the GORT-3 compare to other reading scores?
- What is the relationship between AAE feature production and performances on the GORT-3?

METHOD

Participants

The participants were 65 typically developing African American second through fifth graders who resided in either a midsize central city or an urban-fringe community in the metropolitan Detroit area. These students were recruited from the same school districts that have participated in previous investigations conducted at the University of Michigan using the same recruitment procedures (Craig & Washington, 2000, 2002; Craig, Connor, & Washington, 2003; Craig, Washington, & Thompson-Porter, 1998; Washington & Craig, 1998). Principals in participating schools recruited students for the investigation. Students were included in this study on a continuous enrollment basis during the winter and spring school terms. All of the students were speakers of AAE and produced a minimum of two AAE features during spontaneous oral discourse. Grade, gender, and SES were allowed to vary. Grade was allowed to vary because the reading quotient and standard scores on the GORT-3 control for age.

Forty-five students were from MSES backgrounds and 19 were from low-SES (LSES) backgrounds. The SES of 1 student was unknown. SES was determined from one or both of the following sources: the participants' eligibility or ineligibility to participate in the federally funded free or

reduced-price lunch program and/or the scores on the Hollingshead Four Factor Index of Social Status (Hollingshead, 1975) derived from caregiver questionnaires. This index assigns point scores based on the occupation, years of schooling, marital status, and gender of the child's primary caregiver(s). Point scores ranged from 8 to 66. The point totals corresponded to one of five levels designed to index a family's SES. A female head of household who was assigned 40 points or greater (Level 1 or 2) was placed into the MSES category; a female head of household whose point total was less than 40 (Level, 3, 4, or 5) was assigned the LSES classification. Table 1 lists the distributions of participants by gender, SES, and grade.

Only children who appeared to be typically developing were enrolled in this investigation. The children were judged to be typically developing by their teachers and parents. They had no history of referral to or enrollment in special education services of any type. In addition, each child was administered the Triangles subtest of the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983). Triangles is a matching task that taps a fairly general cognitive skill, is appropriate for children in this age range, and evidences no racial or cultural biases (Kaufman, 1973; Lampley & Rust, 1986; Palmer, Olivarez, Willson, & Fordyce, 1989; Willson, Nolan, Reynolds, & Kamphaus, 1989). All students achieved a scaled score of 7 or more, which is performance within 1 *SD* (-3) of the mean (10), indicating that the participants were within normal limits cognitively. The mean Triangles scaled score for the group was 10.4 ($SD = 2.2$). In addition to typical cognitive skill, the oral vocabulary of the participants appeared to be within normal limits on the Peabody Picture Vocabulary Test—Third Edition (PPVT-III; Dunn & Dunn, 1997). The mean standard score on the PPVT-III was 97.0 ($SD = 11.7$). The higher mean PPVT-III score achieved by participants in the current investigation as compared to a previous study (Washington & Craig, 1999) likely was due to SES differences $F(1, 45) = 5.9, p = .021, \chi^2 = .11$ and schooling effects on vocabulary knowledge.

Table 1. The distribution of the participants by grade, gender, and low (LSES) or middle (MSES) socioeconomic status of the family.

Gender	SES	Grade				Total
		2nd	3rd	4th	5th	
Male	LSES	2	7	0	3	12
	MSES	3	6	4	5	18
	Total	5	13	4	8	30
Female	LSES	1	5	1	0	7
	MSES	7	9	6	5	27
	Total	8	14	7	5	34
Total		13	27	11	13	64*

*The four students who did not produce African American English during oral reading were included in these distributions. The SES of 1 fifth-grade girl was unknown and is not included in this distribution.

Data Collection and Analysis

Administration. The GORT-3, a measure of reading rate, accuracy, fluency, and comprehension, was administered individually to each participant following the instructions provided in the test manual. Both the child and the examiner were audio-recorded with a Marantz PMD430 (Itasca, IL) audio recorder in a quiet room in the child's school using Audio Technica ATH-COM1 (Stow, OH) head-worn microphones.

The GORT-3 is designed for children aged 7;0 (years;months) to 18;11. It consists of 13 passages. Each passage includes one paragraph that is centered on a single topic. Across the test, the passages increase in paragraph length, sentence length, grammatical complexity, and vocabulary difficulty. The test yields raw scores, standard scores, percentiles, and grade-equivalent scores. It provides separate assessments of fluency and comprehension. The fluency assessment consists of a Rate (i.e., the length of time taken to read each passage) and an Accuracy (i.e., number of deviations from print) Score that are combined to form a Passage Score. The Comprehension Score is based on students' responses to multiple-choice questions following each story. The mean for the four subtest components is 10, with a standard deviation of 3. An overall Oral Reading Quotient (ORQ) is derived from the sums of standard scores of the Passage and Comprehension Scores. The ORQ is based on a distribution, with a mean of 100 and a standard deviation of 15.

Students were presented with individual booklets containing all of the GORT-3 passages. They were instructed to read each story aloud as quickly and accurately as possible. The starting passage was selected from the published guidelines. Specifically, second graders began with story one, third and fourth graders with story three, and fifth graders with story five.

Scoring. Following the standardized procedures, individual basals were established separately for the Passage Score and the Comprehension Score. For the Passage Score, an individual basal was achieved when a student attained a score of 9 or greater. For the Comprehension Score, an individual basal was achieved when a student attained a score of 5. An individual ceiling was achieved for the Passage Score when a student attained 2 or fewer points. Similarly, an individual ceiling was established for the Comprehension Score when a student scored 2 or fewer points. For this investigation, basals and ceilings were advantageous because they ensured that each student was reading within his or her own reading skill level.

Because ceilings are established separately on this test for the Passage and Comprehension Scores, it was possible for a child to achieve ceiling on one before the other. In other words, on the same story, a child could achieve a ceiling for the Passage Score (i.e., a score of ≤ 2) and a Comprehension Score that would allow him or her to continue to the next story (i.e., a score of ≥ 3). In this case, the child would read subsequent passages aloud and respond to comprehension questions until a ceiling was achieved for comprehension (i.e., a score of ≤ 2). This is noteworthy because it influenced the number of students for

whom additional passages were available, if needed, when AAE scoring adjustments raised the ceiling.

The oral reading of each passage was timed in seconds. Time was converted to Rate Scores using conversion tables. For Accuracy, number of deviations from print was summed. The sums were converted to Accuracy Scores using conversion tables. The Passage Score was the sum of the Accuracy and Rate Scores.

After students read each passage aloud, multiple-choice questions were posed to assess oral reading comprehension. The questions reportedly tap students' understanding of literal, implicit, critical, and affective information from the paragraphs. The examiner read aloud each question and four possible responses, while the students read along in a student booklet. Correct responses were determined by comparing the child's response to preset answers. The Comprehension Score was the sum of the correct responses.

Each reading passage was scored twice. The first scoring identified any and all reading variations from SAE using the standard scoring procedures in the manual. The second scoring identified those reading variations from SAE that were characterized as AAE phonology, morphosyntax, and combinations (phonology + morphosyntax) using established scoring taxonomies for children (Craig, Thompson, Washington, & Potter, 2003; Washington & Craig, 1994, 2002), and credited the Accuracy Score with any AAE variations. Table 2 presents the phonological, morphosyntactic, and combination AAE scoring taxonomies. The features characterized within these child AAE taxonomies demonstrate some overlap with other dialects of English (Oetting & McDonald, 2001). For example, subject-verb agreement ("when we was about to go to church") is a feature of both AAE and Southern White English (Oetting & McDonald, 2001). Accuracy and Rate Scores could change based on the AAE scoring. Crediting variations from print when they reflected the operation of an AAE feature resulted in higher ceilings for some students, and as a result, fewer variations from print were scored. Lower scores resulting from fewer variations affected the Accuracy Score. Further, students read additional passages, which resulted in extra points being assigned for Rate. Higher ceilings had the potential to result in improved overall scores. Whereas Accuracy and Rate raw scores were combined to generate the Passage Score, this Passage Score changed with the AAE scoring, as well. This was not the case for the Comprehension Score, which had its own basals and ceilings. Additionally, the GORT-3 does not require students to read their responses to this subtest aloud and therefore, the Comprehension Score was not affected by the AAE scoring of the oral reading components of the test.

DDMs were calculated for each reading sample. DDMs control for variable feature production associated with the varying length and number of passages that were read by each participant. In this case, DDMs are the total frequencies of occurrence of AAE features divided by the number of words in the printed passage (Craig & Washington, 2002; Craig et al., 1998; Washington & Craig, 1998). Four DDMs were calculated for each oral reading sample; phonological (PhoDDM), morphosyntactic (MorDDM), combination (phonological + morphosyntactic), and total

Table 2. Phonological (9), morphosyntactic (24), and combination (8) types of child African American English with examples (page 1 of 2).

<i>Definition</i>	<i>Code</i>	<i>Example</i>	
Phonological types			
1. Postvocalic consonant reduction Deletions of consonant singles following vowels	PCR	"mouth" /maʊ/ for /maʊθ/	
2. "g" dropping Substitutions of /n/ for /ŋ/ in final word positions	G	"waiting" /wetɪn/ for /wetɪŋ/	
3. Substitutions for /θ/ and /ð/ /t/ and /d/ substitute for /θ/ and /ð/ in prevocalic positions, /f, t/ and /v/ substitute for /θ/ and /ð/ in intervocalic positions and in postvocalic positions	STH	"this" /dɪs/ for /ðɪs/ "both" /bof/ for /boθ/	birthday" "byfde/ for /byθde/ "with" /wɪt/ for /wɪθ/
4. Devoicing final consonants Voiceless consonants substitute for voiced following the vowel	DFC	/hɪs/ for /hɪz/	
5. Consonant cluster reduction Deletion of phonemes from consonant clusters	CCR	"world" /wɪl/ for /wɪld/	
6. Consonant cluster movement Reversal of phonemes within a cluster, with or without consonant reduplication	CCM	"escape" /ekskep/ for /eskep/	
7. Syllable deletion Reduction of an (unstressed) syllable in a multisyllabic word	SDL	"became" /kem/ for /bikem/	
8. Syllable addition Addition of a syllable to a word, usually as a hypercorrection	SAD	"forests" /forɪstɪsɪz/ for /forɪsts/	
9. Monophthongization of diphthongs Neutralization of diphthong	VOW	"our" /ɑr/ substitutes for /aʊr/	
Morphosyntactic types			
1. Ain't <i>Ain't used as a negative auxiliary in have + not, do + not, are + not, and is + not constructions</i>	AIN	"you <u>ain't</u> know that?"	
2. Appositive pronoun Both a pronoun and a noun, or two pronouns, used to signify the same referent	PRO	"and the other people <u>they</u> wasn't"	
3. Completive <i>done</i> <i>Done</i> is used to emphasize a recently completed action	DON	" <u>done</u> set the fire"	
4. Double marking Multiple agreement markers for regular nouns and verbs, and hypercorrection of irregulars	DMK	"he <u>tries</u> to <u>kills</u> him" "two people <u>felled</u> "	
5. Double copula/auxiliary/modal Two modal auxiliary forms are used in a single clause	MOD	"I'm is the boy"	
6. Existential <i>it</i> <i>It</i> is used in place of <i>there</i> to indicate the existence of a referent without adding meaning	EIT	"I think <u>it's</u> a girl or a boy is yelling"	
7. <i>Fitna/sposeta/bouta</i> Abbreviated forms coding imminent action	FSB	"he <u>fitna</u> be ten" "he <u>bouta</u> fall"	
8. Preterite <i>had</i> <i>Had</i> appears before <i>simple past verbs</i>	HAD	"he flew with a strong stick in his claws while the turtle <u>had</u> held the stick fast in her mouth"	
9. Indefinite article <i>A</i> is used regardless of the vowel context	ART	"one day she met <u>a</u> eagle traveling to a far-away lands across the sea"	
10. Invariant <i>be</i> Infinitival <i>be</i> coding habitual actions or states	IBE	"and they <u>be</u> cold"	
11. Multiple negation Two or more negatives used in a clause	NEG	"it <u>not</u> raining <u>no</u> more"	
12. Regularized reflexive pronoun <i>Hisself, theyself, theirselves</i> replace reflexive pronouns	REF	"bouta fall and trying to hold <u>hisself</u> back up.	
13. Remote past <i>been</i> <i>Been</i> coding action in the remote past	BEN	"I <u>been</u> knew how to swim"	
14. Subject-verb agreement Subjects and verbs differ in marking of number	SVA	"Our cat Mimi like_ to sit on the roof"	
15. Undifferentiated pronoun case Pronoun cases used interchangeably	UPC	" <u>her</u> fell"	
16. Zero article Articles are variably included	ZAR	"this cake is (<u>the</u>) best present of all"	
17. Zero copula/auxiliary Copula and auxiliary forms of the verb <i>to be</i> are variably included	COP	"but she always comes down when it (<u>is</u>) time to eat" "then you'(ll) have to wear the brown ones instead"	
18. Zero <i>-ing</i> Present progressive <i>-ing</i> is variably included	ING	"it was go(<u>ing</u>) to be a good birthday"	

continued on next page

Table 2. Phonological (9), morphosyntactic (24), and combination (8) types of child African American English with examples (page 2 of 2).

Definition	Code	Example
Morphosyntactic types (continued)		
19. Zero modal auxiliary <i>Will, can, do, and have</i> are variably included as modal auxiliaries	AUX	"he might __ been in the car"
20. Zero past tense <i>-ed</i> markers are variably included on regular past verbs and present forms of irregulars are used	PST	"as soon as she open(ed) her mouth, she fall straight into the ocean below"
21. Zero plural <i>-s</i> is variably included to mark number	ZPL	"Father went out to buy some pretty flower_"
22. Zero possessive Possession coded by word order so <i>-s</i> is deleted or the case of possessive pronouns is changed	POS	"The boy'(s) grandmother showed him how to put worms on the hook so they would not come off"
23. Zero preposition Prepositions are variably included	ZPR	"she sits and looks (at) birds"
24. Zero <i>to</i> Infinitival <i>to</i> is variably included	ZTO	"that man right there getting ready _ slip on his one foot"
Combination types		
1. Consonant Cluster Reduction + Zero Past Tense	CCR/PST	"mother kiss(ed) them all goodbye"
2. Consonant Cluster Reduction + Zero Plural	CCR/ZPL	"the children made their bed(s) and dressed"
3. Consonant Cluster Reduction + Subject-Verb Agreement	CCR/SVA	"then she jump(s) on the roof"
4. Postvocalic Consonant Reduction + Zero Auxiliary	PCR/AUX	"I'(ve) lost my blue book"
5. Postvocalic Consonant Reduction + Zero Past	PCR/PST	"the boy's grandmother show(ed) him how to put worms on the hook"
6. Postvocalic Consonant Reduction + Zero Plural	PCR/ZPL	"I can't find my red shoe(s)"
7. Postvocalic Consonant Reduction + Possessive	PCR/POS	"the boy'(s) grandmother"
8. Postvocalic Consonant Reduction + Subject-Verb Agreement	PCR/SVA	"Mimi go(es) up the tall tree by the house"

From "Phonological Features of Child African American English," by H. K. Craig, C. A. Thompson, J. A. Washington, and S. L. Potter, 2003, *Journal of Speech, Language, and Hearing Research*, 46, p. 626-627. Copyright 2003 by the American Speech-Language-Hearing Association. Reprinted with permission.

AAEDDM. For example, total AAEDDM was calculated by dividing the total frequencies of occurrence of AAE features (i.e., phonology, morphosyntax, and combinations) by the number of words in the reading passage.

The ORQs were compared to students' performances on national and state reading achievement tests that were administered by the schools. This comparison evaluated how typical the children's performances on the GORT-3 were relative to their performances on other standardized tests. Standardized scores were available for most of the students from the Iowa Test of Basic Skills (ITBS; Hoover, Dunbar, & Frisbie, 2001), the Metropolitan Achievement Test (MAT, 1993), and the Michigan Educational Assessment Program (MEAP, 1999-2001), a state-mandated achievement test. Both ORQs and standardized reading achievement scores were converted to *z* scores in order to equalize the scales and thus allow comparisons across different distributions.

Reliability

Three independent raters re-scored the GORT-3s of 8 participants. Interrater agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements. Percentage agreements were high across

all scoring systems. Interrater agreements for the presence of a reading variation was 86%, and for distinguishing AAE variations from non-AAE variations was 99%.

RESULTS

Most students ($n = 61, 94\%$) produced AAE features while reading passages from the GORT-3. For one of these students, further analyses were not possible because testing was discontinued at the SAE ceiling. For the remaining 60 students, 1,740 variations from print were produced while reading the GORT-3. AAE features comprised 21% of these variations. Seventy-nine percent of the variations from print were not related to AAE. Two analyses, an independent *t* test and a one-way analysis of variance (ANOVA), were run to examine the effects of SES, gender, and grade on total AAEDDMs. Separate analyses were conducted because the small sample precluded using a three-way ANOVA. The results of *t* tests indicated that total AAEDDMs did not vary systematically by SES, $t(57) = .27, p = .80$ or gender, $t(58) = .68, p = .50$ (see Table 3 for means and standard deviations). Accordingly, in subsequent analyses of DDMs, the data were collapsed across these variables. Results of the

Table 3. Means and standard deviations for dialect density measure (DDM) by gender, socioeconomic status (SES), and grade.

DDM	Gender		SES		Grade			
	Male (n = 27)	Female (n = 33)	LSES (n = 17)	MSES (n = 42)	2nd (n = 11)	3rd (n = 24)	4th (n = 11)	5th (n = 14)
Total								
<i>M</i>	.022	.027	.024	.026	.056*	.019*	.019*	.016*
<i>SD</i>	.023	.036	.024	.034	.059	.016	.009	.013
Morphosyntactic								
<i>M</i>	.002	.004	.001	.004	.007	.002	.003	.002
<i>SD</i>	.005	.006	.002	.006	.009	.005	.003	.002
Phonological								
<i>M</i>	.009	.014	.012	.012	.025*	.010*	.008	.007*
<i>SD</i>	.012	.019	.016	.018	.030	.011	.005	.009
Combination								
<i>M</i>	.011	.010	.010	.010	.024*	.007*	.007*	.007*
<i>SD</i>	.016	.017	.020	.020	.032	.010	.003	.008

* $p \leq .04$

ANOVA revealed significant grade effects, $F(3, 56) = 5.6$, $p = .002$. The second graders produced significantly more AAE features than did the third through fifth graders. Total AAEDDMs decreased from .056 to .016. A small effect size ($\chi^2 = .23$) was associated with these differences. The effect size, which reflects the magnitude of statistical findings, likely was influenced by the large standard deviations for DDM within each grade. It appeared therefore, that although total feature production decreased across grades, grade is not a strong influence. Although by fifth grade it is likely that dialect production will have decreased, there may be considerable variability in the timing of these changes for individual students.

Phonological DDMs and combination DDMs varied systematically by grade, $F(3, 56) = 3.38$, $p = .02$, and $F(3, 56) = 3.91$, $p = .01$, respectively. Small effect sizes were associated with phonology ($\chi^2 = .15$) and combination ($\chi^2 = .17$) DDMs, suggesting that the phonological DDM and combination DDM explained 15% and 17% of the variance attributed to grade. Grade-level differences were evident between second and third grades. Second graders produced DDMs at two to three times the level of third, fourth, and fifth graders, with the exception of fourth graders for phonology. Failure to find differences between the second and fourth graders on phonological DDMs may have been associated with the small number of participants being compared. The sample size for both grade levels was 11 and may not have been large enough to detect significant differences between the two groups of students. For the morphosyntactic DDMs, there were no significant differences by grade, $F(3, 56) = 2.2$, $p = .10$ (see Table 3 for means and standard deviations). Considered together, the data indicate that dialect production rates will be low by fifth grade. A significant decrease in dialect production rates occurred between second and third grade. High levels of variability across students at all grades, however, indicated that there was considerable variability in the timing of the decrease in dialect production rates.

COMPARISONS OF AAE AND SAE SCORINGS OF THE GORT-3

Influence of Subject Status Variables on ORQs

One student was excluded from the initial reading analysis because her SES was unknown. For the remaining 59 students who produced AAE features during reading, ORQs were examined for systematic influences of SES, gender, and grade. The results of multivariate analysis of variance (MANOVA) revealed no statistically significant interaction effects for SES, grade, or gender on the SAE or AAE scoring. Although the GORT-3 standard scores controlled for grade, grade-level comparisons were included in the analysis as a check on the statistical properties of the GORT-3. The ORQs did not vary systematically by grade, SAE-ORQ: $F(3, 45) = .63$, $p = .60$; AAE-ORQ: $F(3, 45) = .82$, $p = .49$. No significant differences were evident on either scoring relative to SES, SAE-ORQ: $F(1, 45) = 0.09$, $p = .77$; AAE-ORQ: $F(1, 45) = 0.12$, $p = .73$, or gender, SAE-ORQ: $F(1, 45) = 1.3$, $p = .27$; AAE-ORQ: $F(1, 45) = 1.01$, $p = .31$. The data were collapsed across these variables in subsequent reading analyses and reported for the 60 students who produced AAE during oral reading. Table 4 presents means and standard deviations for the ORQs by gender, SES, and grade.

Comparisons of ORQs

The mean ORQs on the SAE and AAE scoring were significantly different, *paired t*(59) = 4.4, $p < .001$ (see Table 5). The importance of this statistical difference was evaluated in two ways. First, effect sizes were calculated and found to be moderate ($d = .57$), using the criterion of .30 to .70 (Cohen, 1988). This effect size suggests that the scoring system used does make a difference in the

Table 4. Means and standard deviations of Gray Oral Reading Tests—Third Edition (Wiederholt & Bryant, 1992) standard scores by gender, socioeconomic status (SES), and grade.

Measure	Gender		SES		Grade			
	Male (n = 27)	Female (n = 33)	LSES (n = 17)	MSES (n = 42)	2nd (n = 11)	3rd (n = 24)	4th (n = 11)	5th (n = 14)
SAE scoring								
ORQ								
M	92.1	89.0	89.4	90.4	89.0	92.3	92.7	86.7
SD	19.5	15.2	12.0	19.0	12.4	11.2	20.0	12.4
AAE scoring								
ORQ								
M	93.8	91.4	91.0	92.6	89.4	94.0	97.3	88.4
SD	20.1	14.8	11.7	19.1	13.0	11.1	18.3	26.5

Note. SAE - Standard American English, AAE = African American English, ORQ = Oral Reading Quotient.

Table 5. Standard score means and standard deviations of the students ($N = 60$) on the composite Oral Reading Quotient (ORQ), Rate, Accuracy, Passage, and Comprehension Scores of the Gray Oral Reading Tests—Third Edition (GORT-3);

Wiederholt & Bryant, 1992) when scored for all variations from Standard American English (SAE) and for those variations that credited African American English (AAE). The expected standard score means for the subtests were 10 and for the ORQ were 100. Correlations (r) and regressions (R^2) between the SAE-ORQ and AAE-ORQ and achievement scores appear at the bottom of the table.

Test measure	SAE	AAE
Rate Score (SS)		
M	9.08*	9.23*
SD	3.39	3.32
Accuracy Score (SS)		
M	7.90*	8.93*
SD	3.81	3.72
Passage Score (SS)		
M	8.50*	9.13*
SD	3.46	3.38
Comprehension Score (SS)		
M	8.27	na
SD	3.42	
ORQ (SS)		
M	90.4*	92.5*
SD	17.2	17.3
Achievement (z score)		
r	.69*	.71*
R ²	.48*	.50*

Note. na = not appropriate.
* $p \leq .019$

interpretation of scores. Second, the ORQs were evaluated relative to their potential clinical implications. The ORQ means on the SAE scoring of the GORT-3 were low, 90.4, considerably below the expected standard score mean of

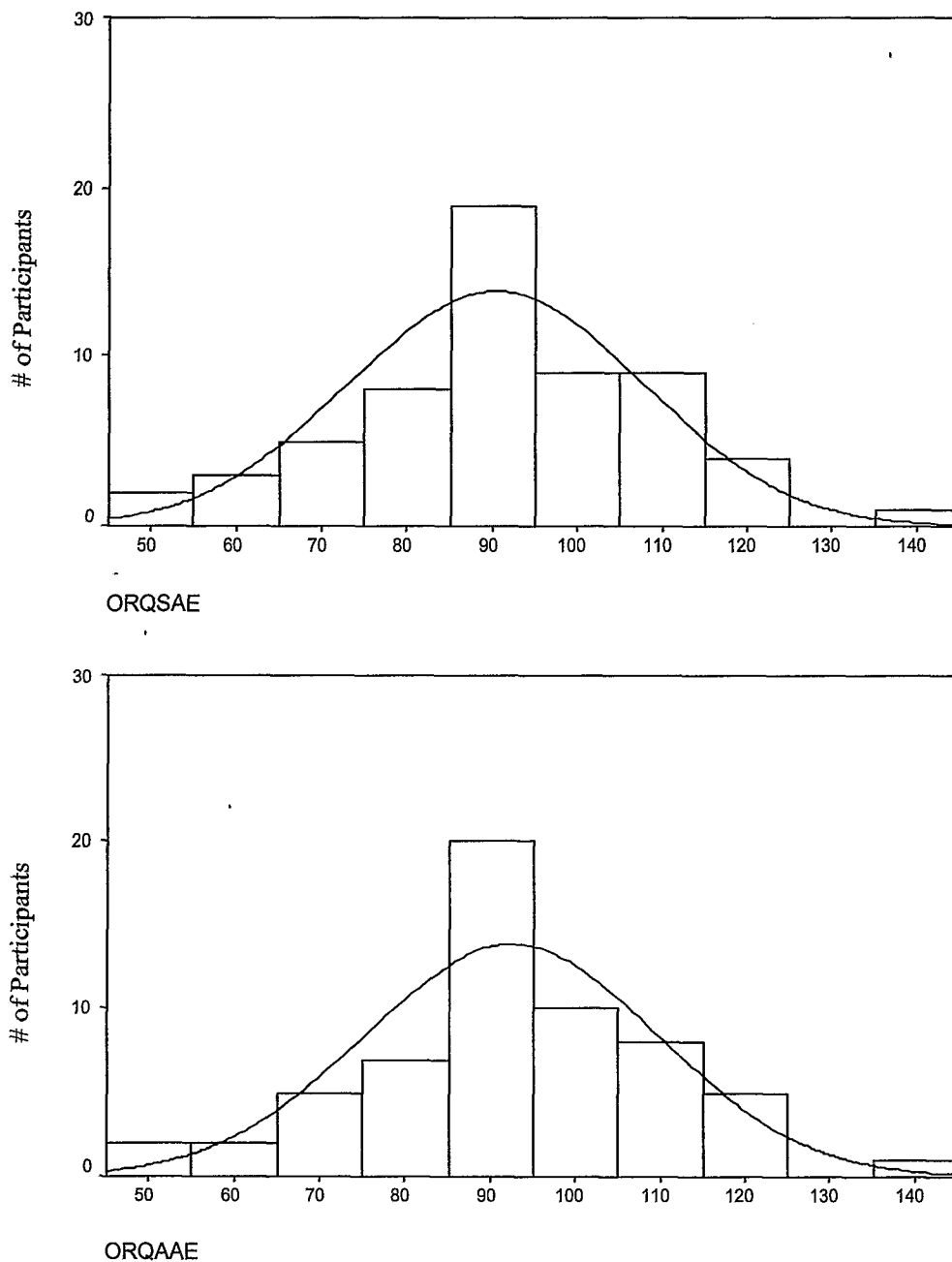
100, and falling just at the boundary between *average* and *below average* on the test (see Table 5). However, the AAE scoring increased the group mean up to 92.5. This score corresponded to a group performance within the *average* performance range on the GORT-3.

The distributions for the SAE and AAE scores across the sample of participants are presented in Figure 1. The ORQs visually approximated a normal distribution in both cases, and this was confirmed with application of the One Sample Kolmogorov-Smirnov test of normality with Lilliefors correction, $K-S(60) = .624, p = .83$ and $K-S(60) = .774, p = .59$, respectively, for the SAE and AAE distributions. On the SAE scoring, the numbers of students falling more than 1 *SD* below the expected standard score mean (i.e., < 85) was 30% of the participant sample. On the AAE scoring, 27% of the students fell more than 1 *SD* below the standard score mean.

Comparisons of Subtest Scores

The Accuracy and Rate Scores, like the overall ORQs, both improved significantly with the AAE scoring, *paired* $t(59) = 6.71, p < .001$, *paired* $t(59) = 2.42, p = .019$, respectively. A moderate effect size ($d = .31$), using the criterion of .30 to .70 (Cohen, 1988), was associated with the difference between the SAE and AAE Accuracy Scores; a large effect size ($d = .87$), using the criterion of .80 or greater (Cohen, 1988), was associated with the difference between the SAE and AAE Rate Scores. The Passage Score was derived from the Accuracy and Rate Scores and so was significant as well, *paired* $t(59) = 5.33, p = .000$. A large effect size ($d = .68$) was associated with the difference between the SAE and AAE Passage Scores. The effect sizes associated with the above findings indicate that the magnitude of these findings was considerable. Comprehension performances were unchanged by the AAE scoring, because the Comprehension Score was independent of the Passage score and was based on student responses to multiple-choice questions (see Table 5 for standard score means and standard deviations).

Figure 1. Distribution of participants' Oral Reading Quotients (ORQs) for the Standard American English (SAE) and African American English (AAE) scorings.



RELATIONSHIP OF PERFORMANCES ON THE GORT-3 TO OTHER READING ACHIEVEMENT TESTS

Group administered standardized reading achievement scores were available for a subset of the participants ($n = 50, 78\%$). In order to permit comparisons across measures, the standardized reading achievement scores and the GORT-3 standard scores were converted to z scores. Both SAE-ORQs and AAE-ORQs evidenced statistically significant,

strong positive relationships to achievement on the independently administered group reading achievement tests, $r = .69, p < .001$; $r = .71, p < .001$ (see Table 5). There were no statistically significant differences between the SAE- and AAE-ORQs and the reading achievement z scores, $\text{paired } t(49) = .480, p = .57$; $\text{paired } t(49) = .061, p = .95$, respectively. Further, SAE-ORQs ($R^2 = .48, p \leq .001$) and AAE-ORQs ($R^2 = .50, p \leq .001$) each predicted half of the variance of the achievement test scores. These findings indicate that performances for these students on the GORT-3, with or without the scoring adjustment, bore

strong predictive relationships to their performances on other standardized tests of reading achievement.

RELATIONSHIP OF DDM TO READING OUTCOMES

Students demonstrated a significant, low, negative correlation between total AAEDDM and Accuracy, $r = -.35$, $p = .006$, and total AAEDDM predicted Accuracy, but at a very low level, $R^2 = .12$, $p = .006$. A significant, low, negative association also was obtained between total AAEDDM and Rate, $r = -.26$, $p = .04$, and again total AAEDDM predicted Rate at a significant, but very low level, $R^2 = .07$, $p = .04$. Total AAEDDM showed no significant relationship to Comprehension, $r = -.21$, $p = .11$. No significant relationships were evident between Accuracy, Rate, or Comprehension and PhoDDM, MorDDM, or combination DDMs.

Considered together, as AAE use increased, Accuracy and Rate decreased, indicating that total feature production impacted these aspects of reading performances, but the influences were not strong. Amounts of AAE produced during reading did not appear to relate to Comprehension.

DISCUSSION

This investigation examined the performances of elementary-grade African American students on the GORT-3. The findings of the present study add some new and useful information to SLPs as they contribute to the planning for literacy acquisition of African American students in their schools. The major findings are as follows:

- African American students produced measurable amounts of AAE during oral reading of SAE text.
- The GORT-3 was appropriate for assessing the oral reading of SAE text by African American students.
- DDMs were related to Accuracy and Rate.

Each of these major findings is discussed below.

Students Produced Measurable Amounts of AAE During Oral Reading of SAE Text

Consistent with prior examinations of the oral reading of African American students (Goodman & Buck, 1973; Harber, 1982; Seymour & Ralabate, 1985), measurable amounts of AAE were produced while reading texts that were written in SAE. Further, nearly all of the participants in the present study produced AAE during oral reading. Twenty-one percent of the variations from print reflected the operation of AAE features. Total AAEDDMs decreased from .056, or one AAE feature for every 18 words for second graders, to .016, or one feature for every 63 words for fifth graders during oral reading. There is no reason to assume that this pattern is absent from silent reading. A core question for future research, therefore, will be to ask

whether this amount of AAE is important to the level of achievement ultimately attained by students.

Interestingly, the total AAEDDM values (.016-.056) were smaller than the MorDDMs (.024-.196) produced by participants in earlier studies conducted at the University of Michigan (Craig & Washington, 2000, 2002; Washington, Craig, & Kushmaul, 1998). Smaller DDMs in the present investigation likely reflected that earlier studies were conducted with preschoolers and kindergartners, who produce higher levels of AAE than elementary-grade students (Craig & Washington, 2003). Additionally, in earlier studies, language samples were collected during free play interactions and picture description, contexts that elicit greater AAE frequencies than oral reading (Thompson, Craig, & Washington, 2003). DDMs during reading of the GORT-3 did not vary systematically by SES or gender. The absence of SES and gender differences likely was the result of the context in which AAE was sampled and schooling effects.

The GORT-3 Was Appropriate for Assessing the Oral Reading of SAE Text

The findings of the present study provide support for use of the GORT-3 when examiners want to determine how a typical African American elementary-grade student will perform when reading texts written in SAE. First, the ORQs were stable regardless of SES and gender. This performance stability further recommends the test for use with African American students and suggests that it will be useful across whole classrooms and a range of communities. The ORQs were also stable across grades. Whereas the construction of the test controls for grade, it is not surprising that the ORQs were comparable across grades. The distributions confirm that the standard scores did indeed control for raw score changes associated with grade for these African American students.

The absence of SES differences on the reading performances of children from LSES and MSES homes in the present investigation is not surprising given that the participants in the study resided in the same community and received similar reading instruction. It is likely that the effects of schooling may have eliminated the effects of SES on students' reading performances. Moreover, SES may not be the most influential variable for explaining performance differences on reading measures during the elementary grades. However, given the small and unequal number of participants included in the LSES category in the present study, further research is warranted to explore this issue more comprehensively.

Second, the statistical differences between the SAE- and AAE-ORQs did not appear to increase the sensitivity or specificity of the GORT-3. In particular, the mean ORQ standard scores for the adjusted and standardized scoring differed by only two points. This measurable difference resulted in a slight change in the classifications of the reading performances of the participants. The SAE-scoring of the GORT-3 placed the participants just at the boundary between *below average* and *average*, and the AAE-scoring placed performances within the *average* range. The minor

increase in the mean standard score and the slight interpretive change suggests that scoring adjustments were not educationally significant and did not appear to be necessary when using this instrument for the purposes described above.

Third, the statistical properties of the performance distributions appeared appropriate. The ORQs were normally distributed, and approximation of the normal curve is important for standardized tests. When participant samples do not distribute normally, one must question the appropriateness of the test for use with the population that the sample represents. Further, in the current investigation, performances on the GORT-3 showed strong associations to other reading tests. Tests that purport to measure the same construct, in this case, reading skills, should evidence high levels of association with each other. Statistical comparisons between both the SAE-ORQs and the AAE-ORQs, and performances on state and national reading tests, revealed strong positive correlations. These findings indicate that students who perform well on the GORT-3 also will perform well on the ITBS, MAT, and MEAP. Indeed, the low values for SAE- and AAE-ORQs obtained in this study are consistent with the national trends for African American students on other tests of reading skill that use SAE texts (Donahue et al., 2001). It is unfortunate that this performance mean for African American students is low. The data from this investigation, in the context of low performance means on other measures, provide additional empirical support to the view that many African American students are having difficulty learning to read SAE texts.

The ORQs predicted approximately half of the variance in the ITBS, MAT, and MEAP scores. To the extent that practitioners want to estimate a student's performance on a state or national test of reading achievement, the GORT-3 is strongly predictive. This information may be useful for SLPs as they assist teachers planning for the literacy development of individual African American students. The GORT-3 offers practitioners a systematic format for the evaluation of variations from print, in terms of both number and type. Further, the performances of groups of students may be informative for formulating larger curricular plans.

The statistical characteristics of the current performances obtained with the GORT-3 recommend its use for assessing the skills of African American students reading SAE texts. This does not mean that the current data support use of the GORT-3 for other interpretive purposes beyond those examined in this study. Most critically, the current data represent performance distributions only and do not address the potential of a student to learn to read or to improve his or her reading skills. It would be unfortunate if practitioners lowered their expectations for African American students by misinterpreting the low ORQ performance means on the GORT-3 as a ceiling on ability.

Although the GORT-3 continues to be widely used in many school settings, a revised version, the Gray Oral Reading Tests—Fourth Edition (GORT-4; Wiederholt & Bryant, 2001) is currently available. The GORT-3 and GORT-4 differ in three important ways. First, one new story was added at the beginning of each form of the

GORT-4. Second, students were given additional time to read stories that appeared at the beginning of the GORT-4. Further, on the GORT-4, the reading variations that were permitted per story were increased. In an effort to evaluate these changes in the GORT-4 for African American students, the performances of 3 readers who scored at the high end of our performance distribution (ORQ range = 115–142) were reexamined. These 3 students were the only ones in our participant sample for whom the stories within their basal and ceiling ranges on both versions of the test were identical regardless of the scoring changes to the GORT-4. Although subtest scores showed slight improvements between test versions, these changes did not affect the overall ORQ scores. It was not possible to make the comparisons between the two test versions for students other than the 3 described above. However, like the higher performing readers, it is possible that the readers who scored at the low end of the performance distribution (ORQ range = ≤ 85) also would not benefit from the GORT-4 changes. Whereas the lower performing students read an average of only 2.6 passages on the GORT-3, raising scores by a few points per story when the number of stories that you can read remains very small would not likely be sufficient to impact overall outcomes.

The data in the current study suggest that there is a trade-off between AAE and fluency (called Passage score in the GORT-3 and Fluency score in the GORT-4), calculated as combined Accuracy and Rate scores on these tests. This hypothesis would predict that increased time and greater allowances for reading variations would positively impact students whose performances fall within the middle range of the performance distribution. The GORT-4 provides an opportunity to probe this potential relationship. A cursory examination of stories that matched between the GORT-3 and the GORT-4 showed improvement in Accuracy and Rate scores with the longer administration times. Although the GORT-3 remains appropriate for AAE-speaking students, the GORT-4 may offer some advantages. In particular, with the benefit of longer administration times, AAE-speaking students may improve their Rate and Accuracy scores at a level sufficient to improve their ORQs. This issue warrants systematic study because of the widespread use of the GORT for research and teaching purposes.

DDMs Were Related to GORT-3 Reading Outcomes

Prior research failed to find important relationships between AAE and reading (Hart et al., 1980; Steffensen et al., 1982). This research was hampered, however, by reliance on adult theoretical models and information gleaned from adult AAE. Accordingly, it is not clear whether the features selected for examination with children in prior investigations were sufficiently sensitive to detect feature/reading linkages. In the current investigation, dialect productions were characterized more globally in terms of dialect density. Every feature produced was included in this analysis rather than a select subset, as has characterized the prior research on this topic. The DDM calculates frequencies of any and

all features in terms of overall word production. The overall DDM evidenced low, negative, statistically significant correlations with Accuracy and Rate and predicted performance on both reading subtests at low, but significant, levels. As a whole, these findings for DDM lend support to the view that studies of the reading acquisition of African American students would benefit from approaches that move beyond the search for specific features, and conceptualize African American readers more broadly as dialect users. Clearly, factors other than dialect production are contributing to reading outcomes as well.

Unlike the predictive relationships between overall DDM and Accuracy and Rate, DDM was not associated with the Comprehension outcomes. This finding is consistent with those earlier studies of AAE feature/reading linkages that found no relationship between AAE features and reading comprehension (e.g. Goodman & Buck, 1973; Harber, 1982). It is beyond the scope of the present investigation to determine whether the lack of association was because AAE, particularly phonological forms, does not influence reading comprehension, or whether DDM is not the best way to detect potential influences on understanding passages. These possibilities should be pursued in future research.

It is important to emphasize for the study as a whole that the findings do not imply that being an AAE speaker is in any way wrong. Before much exposure to schooling, African American students evidence a positive relationship between advanced syntactic and semantic skills and amount of dialect produced (Craig & Washington, 1994, 1995). Further, dialect is self-affirming and culturally positive (Ogbu, 1988).

In conclusion, the present investigation found that use of the GORT-3 was appropriate for examining the performances of elementary-grade African American students reading SAE text. SLPs are in a unique position to influence the reading program and curriculum development for typically developing African American students. Armed with information about AAE, clinicians are equipped to help teachers discern dialectal variations from reading errors. Future research probing the interface between child AAE and reading performances, and the factors supporting dialect-shifting to SAE when learning to read, is clearly warranted. Further, researchers need to determine the role of dialect in teaching reading and ways for educational practice to support the AAE-speaking child learning to read. Intensive inquiry of this type has tremendous potential to improve the literacy outcomes of African American students in positive ways that do not diminish the heritage language of these children.

ACKNOWLEDGMENTS

This work 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, and by the U.S. Department of Education, Office of Educational Research and Improvement, Grant R305T990368. The authors gratefully

acknowledge the cooperation of the students, families, and school personnel participating in this study.

REFERENCES

- American Speech-Language-Hearing Association.** (2001). *Roles and responsibilities of speech-language pathologists with respect to reading and writing in children and adolescents (guidelines)*. Rockville, MD: Author.
- Arnold, K. S., & Reed, L.** (1976). The Grammatic Closure subtest of the ITPA: A comparative study of Black and White children. *Journal of Speech and Hearing Disorders, 41*, 477–485.
- Ball, A. F.** (1994). Language, learning, and linguistic competence of African American children: Torrey revisited. *Linguistics and Education, 7*, 23–46.
- Chall, J. S., Jacobs, V. A., & Baldwin, L. E.** (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Cohen, J.** (1988). *Statistical power analysis of the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Craig, H. K., Connor, C. M., & Washington, J. A.** (2003). Early positive predictors of later reading comprehension for African American students: A preliminary investigation. *Language, Speech, and Hearing Services in Schools, 34*, 31–43.
- Craig, H. K., Thompson, C. A., Washington, J. A., & Potter, S. L.** (2003). Phonological features of child African American English. *Journal of Speech, Language, and Hearing Research, 46*, 623–635.
- Craig, H. K., & Washington, J. A.** (1994). The complex syntax skills of poor, urban, African American preschoolers at school entry. *Language, Speech, and Hearing Services in Schools, 25*, 181–190.
- Craig, H. K., & Washington, J. A.** (1995). African American English and linguistic complexity in preschool discourse: A second look. *Language, Speech, and Hearing Services in Schools, 26*, 87–93.
- Craig, H. K., & Washington, J. A.** (2000). An assessment battery for identifying language impairments in African American children. *Journal of Speech, Language, and Hearing Research, 43*, 366–379.
- Craig, H. K., & Washington, J. A.** (2002). Oral language expectations for African American pre-schoolers and kindergartners. *American Journal of Speech-Language Pathology, 11*, 59–70.
- Craig, H. K., & Washington, J. A.** (2003). *Grade-related changes in the production of African American English*. Manuscript submitted for publication.
- Craig, H. K., & Washington, J. A.** (in press). Language variation and literacy learning. In K. Apel, B. Ehren, E. Silliman, & C. A. Stone (Eds.), *Handbook of language and literacy development and disorders*. New York: Guilford Press.
- Craig, H. K., Washington, J. A., & Thompson-Porter, C.** (1998). Average C-unit lengths in the discourse of African American children from low income, urban homes. *Journal of Speech, Language, and Hearing Research, 41*, 433–444.
- Cunningham, P. M.** (1976-77). Teachers' correction responses to Black-dialect miscues which are non-meaning-changing. *Reading Research Quarterly, 4*, 637–653.
- Delpit, L.** (1995). *Other people's children: Cultural conflict in the classroom*. New York: New Press.

- Dillard, J. L. (1972). *Black English: Its history and usage in the United States*. New York: Random House.
- Donahue, P. L., Finnegan, F. J., Lutkus, A. D., Allen, N. L., & Campbell, J. R. (2001). *The nation's report card: Fourth-grade reading 2000 (NCES 2001-499)*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.
- Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test-III*. Circle Pines, MN: American Guidance Service.
- Federal Interagency Forum on Child and Family Statistics. (2000). *America's children: Key national indicators of well being*. Washington, DC: Government Printing Office.
- Fishback, P. V., & Baskin, J. H. (1991). Narrowing the Black-White gap in child literacy in 1910: The roles of school inputs and family inputs. *The Review of Economics and Statistics*, 73, 725-728.
- Gemake, J. S. (1981). Interference of certain dialect elements with reading comprehension for third graders. *Reading Improvement*, 18, 183-189.
- Goodman, K. S., & Buck, C. (1973). Dialect barriers to reading comprehension revisited. *The Reading Teacher*, 27, 6-12.
- Gray, W. S., & Robinson, H. M. (1967). *Gray Oral Reading Test*. Indianapolis, IN: Bobb-Merrill.
- Harber, J. R. (1977). Influence of presentation dialect and orthographic form on reading performance of Black, inner-city children. *Educational Research Quarterly*, 2(2), 9-16.
- Harber, J. R. (1982). Accepting dialect renderings of extant materials on Black-English speaking children's oral reading scores. *Education and Treatment of Children*, 5, 271-282.
- Hart, J. T., Guthrie, J. T., & Winfield, L. (1980). Black English phonology and learning to read. *Journal of Educational Psychology*, 72, 636-646.
- Hester, E. (1996). Narratives of young African American children. In A. Kamhi, K. Pollock, & J. Harris (Eds.), *Communication development and disorders in African American children* (pp. 227-245). Baltimore: Brookes.
- Hollingshead, A. B. (1975). *Four Factor Index of Social Status*. Unpublished manuscript, Yale University, New Haven, CT.
- Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2001). *Iowa Tests of Basic Skills*. Chicago: Riverside.
- Kaufman, A. (1973). Comparison of the performance of matched groups of Black children and White children on the Wechsler Preschool and Primary Scale of Intelligence. *Journal of Consulting & Clinical Psychology*, 41, 186-191.
- Kaufman, A., & Kaufman, N. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.
- Labov, W. (1998). Co-existent systems in African American English. In S. S. Mufwene, J. R. Rickford, G. Bailey, & J. Baugh (Eds.), *African American English: Structure, history, and use* (pp. 110-153). London: Routledge.
- Labov, W., Baker, B., Bullock, S., Ross, L., & Brown, M. (1998). *A graphemic-phonemic analysis of the reading errors of inner city children*. Unpublished manuscript. Retrieved June 6, 2001, from <http://www.Ling.upenn.edu/~wlabov/Papers/GAREC/GAREC.html>
- Ladson-Billings, G. (1994). *The dreamkeepers: Successful teachers of African American children*. San Francisco: Jossey Bass.
- Lampley, D. A., & Rust, J. O. (1986). Validation of the Kaufman Assessment Battery for Children with a sample of preschool children. *Psychology in the Schools*, 23, 131-137.
- Markham, L. R. (1984). "De dog and de cat": Assisting speakers of Black English as they begin to write. *Young Children*, 39(4), 15-24.
- Melmed, P. J. (1973). Black English phonology: The question of reading interference. In J. L. Laffey & R. W. Shuy (Eds.), *Language differences: Do they interfere?* (pp. 70-85). Newark, DE: International Reading Association.
- Metropolitan Achievement Tests* (7th ed.). (1993). San Antonio, TX: Harcourt Brace.
- Michigan Educational Assessment Program, State Summary Essential Skills Reading Grade 4*. (1999-2001). Lansing, MI: State of Michigan.
- Morgan, M. (1988). More than a mood or an attitude: Discourse and verbal genres in African-American culture. In S. S. Mufwene, J. R. Rickford, G. Bailey, & J. Baugh (Eds.), *African American English: Structure, history, and use* (pp. 251-281). London: Routledge.
- Oetting, J. B., & McDonald, J. L. (2001). Nonmainstream dialect use and specific language impairment. *Journal of Speech, Language, and Hearing Research*, 44, 207-223.
- Ogbu, J. (1988). Cultural diversity and human development. *New Directions for Child Development*, 42, 11.
- Palmer, D. J., Olivarez, A., Willson, V. L., & Fordyce, T. (1989). Ethnicity and language dominance: Influence on the prediction of achievement based on intelligence test scores in non-referred and referred samples. *Learning Disability Quarterly*, 12, 261-274.
- Seymour, H. N., & Ralabate, P. K. (1985). The acquisition of a phonologic feature of Black English. *Journal of Communication Disorders*, 18, 139-148.
- Singham, M. (1998). The canary in the mine: The achievement gap between Black and White students. *Phi Delta Kappan*, 80, 8-15.
- Smitherman, G. (1998). Word from the hood: The lexicon of African-American Vernacular English. In S. S. Mufwene, J. R. Rickford, G. Bailey, & J. Baugh (Eds.), *African American English: Structure, history, and use* (pp. 203-225). London: Routledge.
- Steffensen, M. S., Reynolds, R. E., McClure, E., & Guthrie, L. F. (1982). Black English Vernacular and reading comprehension: A cloze study of third, sixth, and ninth graders. *Journal of Reading Behavior*, XIV(3), 285-298.
- Stockman, I. J. (1996). Phonological development and disorders in African American children. In A. G. Kamhi, K. E. Pollock, & J. L. Harris (Eds.), *Communication development and disorders in African American children* (pp. 117-153). Baltimore: Brookes.
- Taylor, O. L., & Payne, K. T. (1983). Culturally valid testing: A proactive approach. *Topics in Language Disorders*, 3, 8-20.
- Thompson, C., Craig, H. K., & Washington, J. A. (2003). *Variable production of African American English across oracy and literacy contexts*. Manuscript submitted for publication.
- Vaughn-Cooke, F. B. (1986). The challenge of assessing the language of nonmainstream speakers. In O. L. Taylor (Ed.), *Treatment of communication disorders in culturally and linguistically diverse populations* (pp. 23-48). Boston: College-Hill Press.
- Washington, J. A. (2001). Early literacy skills in African-American children: Research considerations. *Learning Disabilities Research and Practice*, 16, 213-221.
- Washington, J. A., & Craig, H. K. (1992). Performances of low-income, African American preschool and kindergarten children

- on the Peabody Picture Vocabulary Test-Revised. *Language, Speech, and Hearing Services in Schools*, 23, 329-333.
- Washington, J. A., & Craig, H. K. (1994). Dialectal forms during discourse of poor, urban, African American preschoolers. *Journal of Speech and Hearing Research*, 37, 816-823.
- Washington, J. A., & Craig, H. K. (1998). Socioeconomic status and gender influences on children's dialectal variations. *Journal of Speech, Language, and Hearing Research*, 41, 618-626.
- Washington, J., & Craig, H. (1999). Performances of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test-III. *Language, Speech, and Hearing Services in Schools*, 30, 75-82.
- Washington, J. A., & Craig, H. K. (2002). Morphosyntactic forms of African American English used by young children and their caregivers. *Applied Psycholinguistics*, 23, 209-231.
- Washington, J. A., Craig, H. K., & Kushmaul, A. J. (1998). Variable use of African American English across two language sampling contexts. *Journal of Speech, Language, and Hearing Research*, 41, 1115-1124.
- Wiederholt, J. L., & Bryant, B. R. (1992). *Gray Oral Reading Tests, Third Edition*. Austin, TX: Pro-Ed.
- Wiederholt, J. L., & Bryant, B. R. (2001). *Gray Oral Reading Tests, Fourth Edition*. Austin, TX: Pro-Ed.
- Wiener, F. D., Lewnau, L. E., & Erway, E. (1983). Measuring language competency in speakers of Black American English. *Journal of Speech and Hearing Disorders*, 48, 76-84.
- Willson, V. L., Nolan, R. F., Reynolds, C. R., & Kamphaus, R. W. (1989). Race and gender effects on item functioning on the Kaufman Assessment Battery for Children. *Journal of School Psychology*, 27, 289-296.
- Wolfram, W. (1994). The phonology of a sociocultural variety: The case of African American Vernacular English. In J. Bernthal & N. Bankston (Eds.), *Child phonology: Characteristics, assessment and intervention with special populations* (pp. 227-244). New York: Thieme.
- Wolfram, W., & Fasold, R. (1974). *The study of social dialects in American English*. Englewood Cliffs, NJ: Prentice-Hall.

Received April 8, 2003

Accepted September 4, 2003

DOI: 10.1044/0161-1461(2004/015)

Contact author: Holly K. Craig, PhD, University Center for the Development of Language and Literacy, University of Michigan, 1111 E. Catherine St., Ann Arbor, MI 48109-2054. E-mail: hkc@umich.edu

A vertical bar on the left side of the page, consisting of a series of horizontal segments in shades of yellow and orange, with a small red diamond at the top.

COPYRIGHT INFORMATION

TITLE: Performance of Elementary-Grade African American
Students on the Gray Oral Reading Tests

SOURCE: Lang Speech Hear Serv Sch 35 no2 Ap 2004

WN: 0409401952004

(C) The American-Speech-Language-Hearing Association is the publisher of this article and holder of the copyright. Further reproduction of this article in violation of copyright is prohibited without the consent of the publisher. To contact the publisher: <http://www.asha.org/>

Copyright 1982-2004 The H.W. Wilson Company. All rights reserved.