A Preliminary Examination of Vocabulary and Word Learning in African American Toddlers From Middle and Low Socioeconomic Status Homes

RaMonda Horton-Ikard  
*University of Tennessee—Knoxville*

Susan Ellis Weismer  
*University of Wisconsin—Madison*

**Purpose:** This study examined the effect of socioeconomic status (SES) on the early lexical performance of African American children.

**Method:** Thirty African American toddlers (30 to 40 months old) from low-SES (*n* = 15) and middle-SES (*n* = 15) backgrounds participated in the study. Their lexical-semantic performance was examined on 2 norm-referenced standardized tests of vocabulary, a measure of lexical diversity (number of different words) derived from language samples, and a fast mapping task that examined novel word learning.

**Results:** Toddlers from low-SES homes performed significantly poorer than those from middle-SES homes on standardized receptive and expressive vocabulary tests and on the number of different words used in spontaneous speech. No significant SES group differences were observed in their ability to learn novel word meanings on a fast mapping task.

**Conclusion:** The influence of socioeconomic background on African American children’s lexical semantic tasks varies with the type of measure used.

**Key Words:** vocabulary, African American toddlers, fast mapping

Linguistic studies of African American (AA) children have focused more often on the forms of language (i.e., grammar and pronunciation) than on its semantic or pragmatic aspects. Examining how AA children use words to convey meaning can be helpful for theory building and clinical work. Theoretically, descriptions about pathways to language development in speakers from culturally and linguistically diverse backgrounds can provide additional evidence for linguistic universals and individual differences. For example, Stockman (1999) reported that AA preschoolers acquire the same lexical and relational meanings as their same-age, White peers. At the same time, AA children are likely to demonstrate some of the unique lexical features that have been identified in adult speakers of African American English (AAE; Green, 2003). Clinically, some research (e.g., Chaney, 1994) has even suggested that early vocabulary knowledge and lexical development play a significant role in the later acquisition of literacy skills.

There are now multiple types of measures for assessing children’s semantic knowledge that can be used for research and clinical purposes. They include three lexical semantic measures of vocabulary or word meaning and its acquisition: (a) norm-referenced-standardized vocabulary tests, (b) lexical diversity measures that are extracted from oral language samples, and (c) word-learning tasks such as fast mapping. Unfortunately, the literature concerning the use of these tools with AA children is scant. These children typically have been excluded from investigations concerned with normal early language development because their inclusion often introduces confounds related to socioeconomic status (SES).

Although the SES of some African Americans has improved since the 1970s, they still tend to be disproportionately represented in welfare and working class groups. They experience longer time periods of poverty than any other ethnic group in the United States (McLoyd, 1997, 1998). The impact of poverty mediates developmental
outcomes through differences in familial material resources, stressful environments, and parental attitudes, values, goals, and knowledge (Garcia-Coll et al., 1996). This outlook suggests that the linguistic behaviors of AA toddlers may be affected by economic as well as cultural factors. Research on how different levels of SES influence developmental outcomes can help to disentangle the effects of poverty and cultural differences. However, SES differences have not been consistently demonstrated across the multiple categories of available measures for assessing vocabulary and word learning. The following review of previous research indicates that there is still a significant gap in the literature concerning the relationship between SES and AA children’s lexical-semantic development.

**Past Vocabulary Studies of AA Children**

**Standardized Vocabulary Tests**

Several studies have focused on the vocabulary performance of AA children using norm-referenced standardized tests such as the Peabody Picture Vocabulary Test—Revised (PPVT–R; Dunn & Dunn, 1981) and PPVT–III (Dunn & Dunn, 1997) and the Expressive Vocabulary Test (EVT; Williams, 1997). Studies of AA children’s performance on standardized tests typically have shown that they score lower than their same-age, White peers (Brookes-Gunn, Klebanov, & Duncan, 1997; J. Campbell, Bell, & Keith, 2001). But many such studies have focused just on AA children from poor and working class families (Brookes-Gunn & Duncan, 1997; Brookes-Gunn, Klebanov, Smith, Duncan, & Lee, 2003; Burchinal, Campbell, Bryant, Wasik, & Ramey, 1997; Champion, Hyter, McCabe, & Bland-Stewart, 2003; Qi, Kaiser, Milan, & Hancock, 2006). As a result, there is little empirical evidence to support the assumption that AA children’s variable scores on standardized tests result from SES as opposed to racial or cultural differences. Even among the studies that have examined the relationship between SES and performance on lexical-semantic measures, the findings have been inconclusive or contradictory.

With respect to receptive vocabulary, AA children’s scores on the PPVT–R and the PPVT–III (Hart & Risley, 1995; Qi et al., 2006) were shown to be vulnerable to SES differences. Hart and Risley’s longitudinal study examined the relationship between linguistic development and SES for 42 children from different racial and class backgrounds using the PPVT–R. Children from upper (n = 13), middle (n = 10), lower (n = 13), and welfare SES groups (n = 6) participated in the study. However, they were reclassified in three distinct groups for the study’s purposes: “professional” (n = 13), “working class” (n = 23), and “welfare” (n = 6). Hart and Risley (1995) determined that the PPVT–R scores of the children from both the working class and welfare groups, at ages 9 and 10, were strongly associated with these same children’s earlier language accomplishments (vocabulary growth rate and use) at 3 years of age. Washington and Craig (1999) concluded that the PPVT–III is appropriate for use with at-risk AA preschoolers, as no significant differences were observed between their mean scores and those reported for this test’s published norms. But the authors also reported that AA children’s performance was differentially influenced by their caregiver’s level of education, one measure of SES.

While receptive vocabulary studies have tended to show SES differences, the evidence of such differences is less conclusive for expressive vocabulary performance. Restrepo et al. (2006) determined that the EVT scores of AA and White preschoolers were significantly influenced by variables related to SES such as maternal level of education, household income, marital status, and the number of children in the family. Maternal education exerted a moderate effect on EVT scores such that children whose parents had a college degree obtained a higher mean standard score (M = 105) compared to mean standard scores ranging from 90 to 97 for those children whose parents were less educated. In contrast, Thomas-Tate, Washington, Craig, and Packard (2006) found no significant SES differences between the EVT scores of AA preschoolers and kindergartners. Taken together, these findings suggest that SES is likely to play a differential role in standardized vocabulary test scores obtained by AA children, although the evidence is less clear for expressive than receptive measures.

**Language Sample Measures**

Language sample analysis has also been used to assess meaningful word use and development. The most frequently used measures include the type-token ratio and the number of different words (NDW). Both are derived from spontaneous speech samples composed minimally of 50 to 100 child utterances. Such measures of lexical diversity are useful for assessing young children’s lexical abilities because they provide information about children’s extant vocabulary use in a natural speech context. Few studies have relied on these tools to describe AA children’s lexical use. Type-token ratio, the ratio of the number of different words divided by the total number of words, was used traditionally to assess lexical diversity (Templin, 1957). But this ratio has been heavily criticized, given the confound between utterance length and the number of word types and tokens used. Longer utterances obviously require the use of more words. In addition, the type-token ratio has been insensitive to changes or differences in chronological age (Klee, 1992; Miller, 1991; Richards & Malvern, 1997).

An alternative to the type-token ratio is the number of different words (NDW). This measure has been shown to differentiate preschoolers with language impairment from their typically developing peers (Conti-Ramsden & Jones, 1997; Watkins, Kelly, Harbers, & Hollis, 1995). NDW also has been shown to be more sensitive than type-token ratio to developmental changes and chronological age (Klee, 1992; Miller, 1991).

The previously described Hart and Risley (1995) study of receptive vocabulary also examined the relationship between SES and measures of lexical diversity. They used “cumulative type-token curves” to map the number of new words used over time by the children from 10 to 36 months of age (p. 164). Primary findings of the investigation indicated that the vocabulary growth of children from the “professional” backgrounds occurred at faster rates than that seen in children from the “working class” backgrounds. Children from “welfare families” demonstrated an average growth of 0.27 or 3 new words (per 100 words spoken),

382 American Journal of Speech-Language Pathology • Vol. 16 • 381–392 • November 2007
compared to 0.60 or 6 new words for children from the “professional families.” Regardless of race, SES influenced the rate of vocabulary growth. Although Hart and Risley concluded that race was not a significant factor in the SES differences observed in vocabulary growth, their sample included just 17 AA children. These children were disproportionately represented in the “working class” and “welfare” groups. Therefore, it may have been premature to conclude that cultural variables related to race do not contribute to differences in vocabulary development.

Dollaghan et al. (1999) also examined the relationship between SES and the diversity of word use in spontaneous speech. Specifically, these investigators examined the relationship between maternal education as the SES measure and NDW in spontaneous speech as the measure of lexical diversity. Spontaneous speech samples were collected from 241 AA and White children who were 3 years old. Specifically, the relationship between maternal level of education and NDW was examined for all participants. Results indicated that maternal educational level influenced NDW such that higher levels of maternal education were related to a larger NDW produced by the children. In addition, those parents with higher education levels tended to be White; just 16% of the AA children’s parents had received a high school education, and 4% of them had a college degree.

Taken together, Hart and Risley (1995) and Dollaghan et al. (1999) indicated that SES is a factor in children’s performance on measures of vocabulary, regardless of race. However, in both studies it was not possible to disentangle SES contributions to linguistic outcomes from other sociocultural variables (i.e., race and ethnicity). To date, very little is known about the vocabulary or diversity of words used by AA children from different SES backgrounds.

**The Fast Mapping Measure**

Novel word-learning tasks, which include fast mapping, also have become useful tools for examining early lexical abilities. A child’s early lexicon is most often acquired through little or no direct teaching. The process that enables children to rapidly increase their vocabulary knowledge involves the mapping of conceptual information received from everyday experiences to numerous linguistic forms (Chapman, 1992). Carey (1978) proposed that children as young as 3 years are able to extract and store information about new lexical items with a minimum amount of exposure and experience with an unfamiliar word. Dollaghan (1985) found that children between 2 and 5 years of age were able to fast map novel words to novel referents using a puppet play activity. Rice, Buhr, and Nemeth (1990) observed that in the course of two 6–8 min television viewing paradigms, typically developing children were capable of comprehending new words and adding them to their lexicon. The ability to accomplish these types of tasks in everyday contexts has been referred to as quick incidental learning by Rice et al. (1990).

To date, just one study appears to have examined AA children’s response to a fast mapping task (Johnson, 2001). Johnson’s participants included 60 AAE and Standard American English (SAE) speaking 4–6-year-old children from low- to low-middle-SES communities. In particular, she examined the relationship between the scores on the PPVT–III and a novel word comprehension task. Participants were exposed to picture sequences in which they guessed the meaning of a novel verb based on its argument structure. Johnson reported that AAE participants performed significantly worse on the PPVT–III than did their SAE peers. But no significant AAE and SAE group differences were observed on the novel word comprehension task. Johnson concluded that dialect did not influence the relationship between PPVT–III scores and the novel verb comprehension task.

Although Johnson’s (2001) study provided a new and innovative way of examining language performance in AA children, it did not aim to test any hypotheses about the relationship between SES and fast mapping performances. It is unknown whether the cultural differences mediated by socioeconomic privilege will operate in the same fashion. Thus, if processing tasks are less susceptible than norm-referenced tests to prior experiential effects related to race and culture, then we could expect few performance differences for children from diverse economic backgrounds.

**Motivation for the Present Study**

The above studies, taken together, indicate that SES may affect AA children’s performance on standardized receptive and expressive vocabulary tests as well as the diversity of words used in spontaneous speech, although the evidence is far from conclusive. It is unknown whether their performance on word-learning tasks such as fast mapping is susceptible to SES differences because such differences have never been explored. Furthermore, no single study appears to have examined SES differences among the same group of AA children across the broad spectrum of lexical-semantic measures available for observation. But there is reason to do such research because the various types of measures described above yielded different types of information about lexical-semantic knowledge. Stockman (2000) pointed out that standardized vocabulary tests such as the widely used PPVT–III are designed to measure existing vocabulary knowledge. As such, they are experience-dependent measures in the sense that they assess the words that children typically have already learned from their life experiences. Similarly, a measure of lexical diversity based on spontaneous speech (e.g., the NDW) is experience-dependent. The experience-dependent measures can be influenced by differences in linguistic input related to family SES and maternal level of education.

In contrast, lexical processing tasks such as fast mapping are arguably less dependent on prior experience. According to deVilliers (2004), fast mapping tasks focus on word processing skills that do not rely on existing knowledge. Therefore, they limit the opportunity for performance to be influenced by cultural and linguistic biases. The expectation is that groups may differ less on experience-independent than dependent measures. This assumption is supported by previous studies in which both types of measures were used to compare groups that differed in other ways besides SES. Such was the case for the AAE and SAE speakers compared in Johnson’s (2001) fast mapping task described above.
The same point was made by another study involving the Non-
Word Repetition Task (T. Campbell, Dollaghan, Needleman, & Janosky, 1997) in which several minority racial groups of children were compared to a group of White children. Campbell et al. revealed that the performances of minority children, inclusive of AAE speakers, were comparable to their same-age White peers on word processing tasks, which assessed their ability to learn new information. But these same minority children scored below their White peers on standardized vocabulary and language measures that were presumed to be experience-dependent.

So far, the same investigative paradigm has not been used to compare AA children from different SES groups. But much might be learned from determining whether low- and middle-SES groups differ on all categories of lexical semantic tasks. The literature review showed that AA children from low-SES homes tend to score lower than their same-age peers from middle-SES homes on standardized vocabulary test norms. However, there is not enough information on the performance of AA children from middle-SES homes to determine if they too tend to score below the norms on standardized vocabulary tests. Consequently, it can be difficult to discriminate AA children’s low scores that may be due to normal cultural and linguistic differences from those due to genuine language impairment. In addition, simply observing AA children’s performance on a single task does not provide us with insight concerning the full scope of their lexical-semantic knowledge. For example, reliance on just norm-referenced and language sample measures of lexical knowledge may reveal something about the words that they already have learned from experience, but not their potential to learn new words, as might be revealed by word processing tasks such as fast mapping. Given the previous research cited above, typical learners from low- and middle-SES groups may differ less on novel word-learning tasks such as fast mapping than on the standardized vocabulary and lexical diversity measures of the extant vocabulary, which are so often used for research and clinical evaluations. If so, then there ought to be practical consequences of showing whether a differential performance effect exists.

The use of multiple types of lexical assessment procedures may allow one to differentiate low performance on standardized vocabulary and lexical diversity measures due to cultural or experiential biases from the low performance due to impaired processing constraints on language learning potential. Children with impaired learning potential ought to perform poorly on novel word-learning tasks, but those without such impairment should not, even if they score lower than expected on norm-referenced and language sample measures of vocabulary. Research designed to test such hypotheses requires the same low- and middle-SES groups of AA children to be assessed and compared on a variety of tasks that include both experience-dependent and less dependent lexical-semantic tasks.

**Purpose of the Current Study**

The purpose of the current investigation was to examine the vocabulary and word-learning abilities of AA children from two different SES backgrounds on three types of measures examining lexical-semantic development. The following questions guided this study:

1. How do typically developing AA toddlers from low- and middle-SES groups perform on commonly used lexical-semantic measures (PPVT–III, EVT, NDW, and a fast mapping task)?

2. Do SES differences vary with the type of lexical-semantic measure?

**Method**

**Participants**

The participants were 30 typically developing AA toddlers (age 30 to 40 months) from a midsize southeastern city. They were recruited via newspaper announcements, flyers, and letters circulated to daycare providers that summarized the purpose and selection criteria of the study. Toddlers whose parents signed consent forms and met the criteria for typically developing status were the participants for this investigation. They were judged to be typically developing if there was no history of clinical referral or parental concern about their motor, cognitive, or linguistic development.

**Developmental Status**

All the children passed a bilateral hearing screening as assessed by distortion product otoacoustic emissions (Joint Committee on Infant Hearing, 2000). Each child scored less than 1.25 SDs from the mean on the Battelle Developmental Inventory (BDI; Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1988) to ensure that typically developing status reported by parents was accurate. Only one child scored below 85 (the cutoff score for 1 SD below the mean), and that child’s BDI score was an 84. All of the children were in the range of normal language performance as identified by the total language score of the Preschool Language Scale, Fourth Edition (PLS–4; Zimmerman, Steiner, & Pond, 2002). No child scored below 85 on the PLS–4. Every child obtained a mean length of utterance in words (MLU-W) score of at least 2.69, which is within age-appropriate norms when compared to Miller and Chapman (1981).

**SES Groups**

Children were assigned to either the low- or middle-SES group based on the Four Factor Index of Social Status (Hollingshead, 1975). It provides a score based on the parental level of education, occupation, gender, and marital status. The Four Factor Index scores can range from 8 to 66 points. The children whose parents scored at 30 or above composed the middle-SES group in this study, and those whose parents scored at or below 29 composed the low-SES group. Information used to assess SES was obtained from a background questionnaire completed by a child’s parent or guardian. It elicited information about family demographics and the child’s developmental and health history. The mean index scores for each SES group are reported in Table 1.
The two SES groups were matched as closely as possible for chronological age. Table 2 identifies the mean ages, BDI scores, PLS–4 scores, and MLU-W for each SES group. Independent-samples t tests were performed to determine whether any of these factors were significantly different for the two groups. There was no significant difference between the two groups in chronological age. BDI mean score for toddlers from the middle-SES group was significantly higher than their low-SES peer group, t(28) = −0.521, p = .607, d = 0.2 (a small effect size). The BDI mean score for toddlers from the middle-SES group was significantly higher than that of the low-SES group, t(28) = −2.07, p = .048, d = 0.78. This difference was associated with a large effect size. The PLS–4 mean score for the middle-SES group was also significantly higher than that of the low-SES group, t(28) = −3.83, p = .001, d = 1.45. This difference was also associated with a large effect size. However, no significant differences were observed between the two SES groups in their language levels, as measured by MLU-W, t(28) = 0.534, p = .597, d = 0.19 (a small effect size).

### Data Collection Procedures

Data were collected across two sessions. During the first session, each child’s hearing was screened, and the background questionnaire, BDI, and PPVT–III (Form A; Dunn & Dunn, 1997) were administered. At the second session, the fast mapping task (Ellis Weismer & Evans, 2002) and the EVT (Williams, 1997) were administered.

In addition, a spontaneous language sample was obtained at the end of the session. The standardized tests and the fast mapping task were administered by the first author, who is African American, and a White, female graduate student majoring in speech-language pathology. The first author collected all of the language samples.

### Norm-Referenced Standardized Vocabulary Tests

The PPVT–III, which is a commonly used vocabulary test, measures receptive vocabulary knowledge using a picture recognition (pointing) format. The EVT is a test of expressive vocabulary knowledge requiring picture labeling and synonym tasks. Both tests are appropriate to use with children as young as 30 months of age. The PPVT–III and the EVT were standardized on populations that included AA children in proportion to their representation in the general population. For these reasons, the PPVT–III and EVT have been considered appropriate to use with AA children. Both tests were administered according to the guidelines provided in the testing manual.

### Language Sample Measure of Lexical Diversity

A 20-min spontaneous language sample was collected from each participant during a semistructured free-play interaction with the examiner. The materials used for the collection of the samples were a Fisher Price Family Home Set, figures, and accessories. Samples were audio-recorded using a Marantz audio-recorder. The examiner wore a lapel microphone, and a PZM microphone was also placed at floor level near the child.

Language samples were orthographically transcribed from the audiotapes using Systematic Analysis of Language Transcripts (SALT; Miller & Chapman, 2000). Utterance segmentation was based on standard SALT procedures. They define an utterance as a verbal production that typically is separated by pauses and/or intonation cues and includes no more than two independent clauses. The Standard Measures analysis provided in SALT was used to extract the NDW measure for a 100-utterance analysis set once the following utterances were excluded: abandoned and interrupted utterances, unintelligible utterances, nonverbal utterances, imitations, and yes/no responses to questions. Four participants fell short of 100 utterances once these specific utterances were excluded, so their analyses were based on the entire sample obtained, which ranged from 82 to 97 utterances.

### Fast Mapping Task

The fast mapping task used in this study was developed by Ellis Weismer and colleagues for a longitudinal study of language development in late talkers and typically developing toddlers (Ellis Weismer & Evans, 2002; Ellis Weismer et al., 2001). Previous studies have used story-based tasks and extended word-learning trials to examine fast mapping and other word-learning tasks in children who were generally older than those observed in the present study.
(Burton & Watkins, 2005; Gray, 2004). The current study focused on initial fast mapping by toddlers. Given their ages, the fast mapping task needed to be an activity that would keep the toddlers engaged for a certain period of time in addition to being age appropriate.

A puppet play activity that involved packing and unpacking a picnic basket lunch was piloted for this study. Pilot observations revealed that this task consistently engaged the toddlers. They also revealed that the type of puppet play activity described below and the number of trials within each phase of the task were developmentally appropriate for toddlers (i.e., resulted in neither floor nor ceiling effects).

Stimuli. Fast mapping during the puppet activity was examined using two novel word labels, one for each of the two unfamiliar objects, along with two familiar labels and objects. Filler trials with familiar objects and labels were included within the task to ensure that the toddlers were consistently engaged in the task. However, only the data related to the mapping of novel words and objects were analyzed for this study. The familiar words, apple and cookie, referred to toy replicas of an apple and a cookie. The unfamiliar or novel words, koob and tade, referred to unusual objects that were predetermined to have low codability by adult judges during an earlier investigation of novel word learning (Ellis Weismer & Hesketh, 1996). The two novel words, koob and tade, also were chosen because they were composed of early developing sounds that are typically present in the phonemic repertoire of 30-month-old toddlers (Stoel-Gammon, 1991).

Testing procedures. During the puppet play activity, the examiner showed the children a puppet and instructed them to help the puppet pack a picnic with all types of things to eat, including some “silly things.” The examiner removed the four items from a bag before engaging the child in the exposure phase of learning. During the exposure phase, the examiner handed the child one of the four items and spoke the name of the object in the carrier phrase, “Here’s a ______. Put it in the basket.” This carrier phrase was repeated until all four of the items (novel and familiar) were placed in the basket. After each novel and familiar word had been labeled, the examiner proceeded to the testing phase, which revealed what a child had retained or learned. The testing phase consisted of a production (expressive) probe phase followed in order by a comprehension (receptive) probe phase. During the production probe phase, the examiner suggested that she and the child look at what was in the puppet’s lunch. The examiner then picked up one of the four objects from the picnic basket and asked, “What’s this?” The named or targeted item was then returned to the picnic basket, and the examiner proceeded to elicit responses in the same manner for the other three objects.

The comprehension probe phase followed the production probes. For the comprehension probe, the examiner said, “Let’s put some more things in too” and added two novel object foils to the picnic basket. In other words, the comprehension task involved the use of six items (the two novel objects, the two familiar objects, and the two foil objects). The foils were also unusual objects with low codability. Their inclusion among the stimuli helped to reduce a child’s probability of selecting the correct item by chance. The examiner pretended that the puppet “whispered” into her ear. Then the examiner said, “He wants to taste them. Can you get the ______?” Once the child responded by choosing an object to feed the puppet, the examiner repeated the carrier phrase for the remaining three target words in the same manner.

After the comprehension phase was completed, the examiner pretended that the puppet decided that he did not want to go on a picnic but then he told the examiner that he had changed his mind again. The examiner said, “Okay, we’ll pack the lunch again.” The entire procedure was then repeated when two other puppets convinced the first puppet to go on another picnic. This procedure provided the opportunity for three exposure trials and response probes for each task, that is, the production and comprehension tasks. The order of presentation for the novel words was counterbalanced across children. Participants did not receive feedback regarding the accuracy of their responses.

Scoring. A graduate assistant, who was blind to the study’s research goals, scored the fast mapping task online. On the comprehension probes, participants received a score of 1 for the correct selection of the targeted object and a score of 0 for an incorrect selection. Productions were judged to be correct if the participants accurately produced two out of the three speech sounds in the appropriate sequence. Audio-recording of the entire task occurred for the purposes of improving scoring accuracy and assessing reliability of the production probes. Correct responses for production task were also given a 1, and incorrect responses were scored as 0. Given the three response opportunities for each word, the accuracy scores for each novel word (koob and tade) could range from 0 to 3. The overall possible score ranged from 0 to 6 for each task. Comprehension and production accuracy scores were each computed as the proportion of accurate or correct responses for the novel words observed across the six probes or opportunities to respond. For example, a child who received 2 out of 3 points for the correct production of koob and 2 out of 3 points for the correct production of tade would have an overall production score of 4 out of 6 (0.67) for novel word targets. Similarly, for the comprehension task, if a child received 1 out of 3 points for the correct selection of the novel object koob and 1 out of 3 points for the correct selection of tade, the child would have earned an overall comprehension score of 2 out of 6 (0.33).

Reliability

Language Transcription Reliability

Ten of the 30 language transcripts were randomly selected for reliability scoring (33%). Half of the selected transcripts included samples from the low-SES group, and the other half were from the middle-SES group. Once the samples were orthographically transcribed, a second person listened to and reviewed them in order to correct errors and note disagreements about the orthographic transcription of utterances with the first transcriber. The percentages of interrater agreement for utterance segmentation and morpheme-by-morpheme transcription were 94% (857/908) and 90% (3026/3359), respectively.
**Interrater Agreement for Fast Mapping Scores**

Although the first author administered the fast mapping task, a graduate student in communicative disorders, who was blind to the purpose of the study, completed the scoring. The student had course work as well as research and clinical experiences with phonetically transcribing the speech of toddlers and preschoolers. The audiotaped responses of the children also were scored by a second person whose background was comparable to the first rater. Recall that the production probes were scored as correct if the child produced 2 out of the 3 sounds accurately in the appropriate sequence. The production data for half of the participants were examined for interrater agreement in judging correct and incorrect responses and the phoneme-by-phoneme analysis. Interrater agreement was 91% (82/90 agreements) for score accuracy and 88% (713/810) for the phoneme-by-phoneme analysis.

**Data Analysis**

Independent-samples \( t \) tests (two-tailed) were used to determine whether there were significant differences between the two SES groups on the norm-referenced and language sample measures. In addition, one-sample \( t \) tests (one-tailed) were used to determine whether each group’s mean score differed significantly from those of the normative sample used to standardize the PPVT–III and the EVT. A repeated measures mixed-model analysis of variance (ANOVA) was used to examine between- and within-group differences for the word-learning/fast mapping task. Cohen’s \( d \) was used to determine effect sizes for all \( t \) tests, and eta squared was used for the \( F \) test in the ANOVA (Tabachnick & Fidell, 1996).

**Results**

**Norm-Referenced Standardized Measures**

**Sample Mean Comparisons**

Table 3 displays the mean score, standard deviation, median, and range for each SES group. A statistically significant SES group difference was observed for the PPVT–III and the EVT. Both significant results were associated with large effect sizes. With respect to receptive vocabulary, Table 3 shows that the mean scores were higher for children in the middle- than the low-SES group on the PPVT–III, \( t(28) = -3.10, p = .004, d = 1.17 \). The mean standard score for the children from the low- and middle-SES groups corresponded, respectively, to the 30th and 57th percentiles for the norms reported in the PPVT–III manual. With respect to expressive vocabulary, Table 3 shows that the mean scores also were significantly higher for children in the middle-SES group, \( t(28) = -3.920, p = .001, d = 1.48 \). The mean standard scores corresponded to the 30th and 58th percentiles for the low- and middle-SES groups, respectively.

**Table 3. Descriptive statistics for the norm-referenced standardized vocabulary measures by SES group.**

<table>
<thead>
<tr>
<th>Tests</th>
<th>Low</th>
<th>Middle</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT–III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M )</td>
<td>92.73</td>
<td>104.00</td>
</tr>
<tr>
<td>( SD )</td>
<td>7.24</td>
<td>12.05</td>
</tr>
<tr>
<td>( Mdn )</td>
<td>93.00</td>
<td>106.00</td>
</tr>
<tr>
<td>Range</td>
<td>80.00–103.00</td>
<td>79.00–119.00</td>
</tr>
<tr>
<td>EVT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M )</td>
<td>91.93</td>
<td>102.60</td>
</tr>
<tr>
<td>( SD )</td>
<td>5.36</td>
<td>9.06</td>
</tr>
<tr>
<td>( Mdn )</td>
<td>92.00</td>
<td>104.00</td>
</tr>
<tr>
<td>Range</td>
<td>79.00–99.00</td>
<td>85.00–115.00</td>
</tr>
</tbody>
</table>


**Sample Mean Comparisons to Normative Test Scores**

The observed mean scores for the PPVT–III and the EVT were each compared to the normative mean scores of their respective 36-month-old age groups using a one-sample \( t \) test.

The PPVT–III. Our observed mean PPVT–III score of 92.73 (SD = 7.24) for the low-SES group was significantly below the mean score of 100 (SD = 15) for this test’s normative group, \( t(14) = -3.884, p = .002, d = 0.51 \). A moderate effect size was observed. However, no significant difference was observed between the middle-SES group and the PPVT–III’s normative sample, \( t(14) = 1.285, p = .220, d = 0.28 \), and a small effect size was associated with this difference.

Further examination of the individual data revealed that only 2 of the children in the low-SES group and 1 child in the middle-SES group scored more than 1 SD below PPVT–III’s mean normative score of 100. Given the PPVT–III’s standard error of measurement (4.7), these three scores, ranging from 79 to 84, were all less than 1.5 SDs below the mean, a frequent criterion for clinical referral. No children in either SES group scored more than 1 SD above the mean of the normative sample.

The EVT. A one-sample \( t \) test revealed that on the EVT, the mean score for the low-SES group was significantly below the mean score of 100 for this test’s 36-month-old age group, \( t(14) = -5.824, p = .001, d = 0.57 \), a medium effect size. But the same analysis for the middle-SES group yielded no statistically significant differences, \( t(14) = -1.11, p = .286, d = 0.18 \), and a small effect size was obtained. Examination of individual scores revealed that all children in the low-SES group scored within 1 SD of the normative sample’s mean score, except for 1 child whose standard score was 79. Such an outlier score was within the range of expected normal score variability for this test, given its standard error of measurement, which was also 4.7. None of the children in either of the two SES groups scored more than 1 SD above the EVT’s standardization sample mean.

**Language Sample Measure: NDW**

Table 4 displays the mean NDW along with its standard deviation and range for each SES group. The independent-samples \( t \) test revealed significant SES group differences. Toddlers from the middle-SES group used a significantly larger NDW than their low-SES peers, \( t(28) = -2.574, \)
and on a language sample measure of lexical diversity (NDW) in a spontaneous language sample. SES did not significantly affect the same children’s overall performance on the fast mapping task. This study may be the first one to systematically examine SES in relationship to a range of lexical assessment procedures that include a novel word-learning task. Its focus on SES differences also reflects the increasing interest in describing the heterogeneity within the AA population as well as how it differs from other racial and ethnic groups, which has been the traditional focus. In the following section, this study’s findings are discussed in terms of their similarities and differences with previous investigations and their implications for future research and clinical practices.

Discussion

The purpose of this investigation was to examine the vocabulary and word-learning abilities of AA children from low- and middle-SES backgrounds. The results revealed that SES was a factor that influenced toddler performances on norm-referenced vocabulary measures (PPVT–III and EVT) and on a language sample measure of lexical diversity (NDW) in a spontaneous language sample. SES did not significantly affect the same children’s overall performance on the fast mapping task. This study may be the first one to systematically examine SES in relationship to a range of lexical assessment procedures that include a novel word-learning task. Its focus on SES differences also reflects the increasing interest in describing the heterogeneity within the AA population as well as how it differs from other racial and ethnic groups, which has been the traditional focus. In the following section, this study’s findings are discussed in terms of their similarities and differences with previous investigations and their implications for future research and clinical practices.

Fast Mapping Task Measure

Table 5 displays the SES group means for the comprehension (receptive) and production (expressive) accuracy rates on the fast mapping task. A repeated measures mixed-model ANOVA revealed no significant between-subject effect for SES, $F(1, 28) = 0.784, p = .383, \eta^2 = .027$. But there was a significant within-subject effect for type of probe modality used (comprehension vs. production), $F(1, 28) = 15.667, p = .001, \eta^2 = .359$. The toddlers exhibited significantly higher scores on the comprehension than the expressive or production phase of the fast mapping task. The within-subject factor of probe type accounted for a moderate amount (nearly 36%) of the variance in score accuracy.

The interaction for SES and probe modality was not significant, $F(1, 28) = 0.715, p = .405, \eta^2 = .025$. The absence of a significant interaction effect indicated that the comprehension modality bias was applicable to both the middle- and low-SES groups. SES accounted for less than 3% of the variance in scores, which is considered a small effect. Overall then, the data supported comparable performances for the low- and middle-SES groups on the fast mapping tasks.

Consistency of Findings With Past Research: SES Group Effects

Experience-Dependent Measures

The observation of SES group differences on norm-referenced and language sample measures of lexical knowledge was consistent with the outcomes of earlier studies that have documented less extensive vocabulary knowledge among children from low-SES backgrounds compared with their middle-SES peers (Hart & Risley, 1995; Qi et al., 2006). In the current study, toddlers in the low-SES group had a mean standard score that was 10 points below that of their normative age groups on both the PPVT–III and the EVT. In contrast, toddlers from the middle-SES homes performed at levels that were comparable to the normative groups for the same two tests. The depressed standardized test scores of the low-SES group, which ranged from 79 to 119, were congruent with the findings of other studies in which the mean standard scores ranged from 87 to 96 for low-SES preschoolers and school-age children (Champion et al., 2003; Restrepo et al., 2006; Washington & Craig, 1999). Qi et al. (2006) found that maternal education influenced PPVT–III scores, such that children whose parents had a bachelor’s degree scored 11 points higher than those children whose parents had no high school diploma.

Restrepo et al. (2006) observed that AA children’s performance was less depressed on the EVT, an expressive vocabulary measure, than it was on the PPVT–III, a receptive vocabulary measure. Their 4–5-year-old AA children from low-income homes performed below the mean normative score but within 1 SD of reported EVT norms. The same outcome was obtained in the current study for the

<table>
<thead>
<tr>
<th>Probe type</th>
<th>Score distribution</th>
<th>SES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Middle</td>
</tr>
<tr>
<td>Comprehension accuracy</td>
<td>$M$</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0.00–1.00</td>
</tr>
<tr>
<td>Production accuracy</td>
<td>$M$</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0.00–1.00</td>
</tr>
</tbody>
</table>

TABLE 4. Mean values for number of different words by SES group.

<table>
<thead>
<tr>
<th>Score distribution</th>
<th>SES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>$M$</td>
<td>113.13</td>
</tr>
<tr>
<td>$SD$</td>
<td>16.55</td>
</tr>
<tr>
<td>Range</td>
<td>79–136</td>
</tr>
</tbody>
</table>

$p = .016, d = 0.97$, a large effect size. In addition, the minimum NDW used by the middle-SES group was higher than that for the low-SES group.
low-SES group. Thus, the current study provided additional evidence that parental education may play a role in the stan-
dardized test scores of AA children. Recall that the edu-
cation scores of the parents of toddlers from middle-SES backgrounds were higher than their peers from low-SES backgrounds.

Caregiver’s educational level appeared to influence not only norm-referenced test scores but also the language sample measure (NDW) used in the current study. Its findings were consistent with those reported by Dollaghan et al. (1999) in revealing that the level of caregiver education was related to the NDW in children’s spontaneous speech.

How can we account for the SES effects on experience-
dependent measures? Some researchers have suggested that the way in which parents structure their children’s language-learning environment influences the acquisition and rate of language development (Arriaga, Fenson, Cronan, & Pethick, 1998; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002). Hoff (2003) suggested that the structuring of the language environment is especially vulnerable to SES-related differences. In addition, Hoff (2003), like Hart and Risley (1995), suggested that SES-related differences in the amount of vocabulary and diversity of vocabulary are due to the nature of child-directed speech provided by parents and caregivers, whether assessed either by norm-referenced or language sample measures. For example, mothers from high-SES backgrounds tend to use more conversation-eliciting questions, more complex syntax, and a greater diversity of vocabulary when talking to their toddlers compared with those mothers from low-SES backgrounds (Hoff, Laursen, & Tardif, 2002; Hoff-Ginsberg, 1998). These investigatory suggestions that it is because parents from high-SES homes talk to their children more, read to their children more, and use language for a greater variety of functions that their children’s vocabulary is larger and grows at a more rapid rate than their lower SES peers.

Word-Learning or Fast Mapping Tasks
The findings from the fast mapping task used in the current investigation revealed no significant SES group differences when simply examining accuracy scores for the comprehension and production of novel lexical targets. This observation reflected the possibility that performance on the fast mapping task depended less on a toddler’s existing knowledge than did the norm-referenced vocabulary tests and the NDW measure. T. Campbell et al. (1997) argued that this type of task is considered “processing dependent.” That is, it required the toddlers to rely less on existing vocabulary knowledge and more on psycholinguistic processing abilities. Such skills tap general cognitive operations (i.e., memory, recall, and retrieval abilities) that make it possible for all children to manipulate and generate language. Consequently, children from the low-SES group may have had less existing vocabulary knowledge than did their middle-SES peers because of their prior language experiences, but their differences were minimized when provided with the same type of linguistic input for learning novel words. Recall that T. Campbell et al. (1997) used another type of processing task, the nonword repetition test, to reveal no racial or minority ethnic group differences. The current study is the first to provide evidence that such tasks might not be vulnerable to SES differences.

Consistency of Results With Past Research
Task Effects
The findings from the current investigation also offer some additional insight into the receptive/expressive bias frequently observed in language performances. This is because the children performed on both receptive (PPVT–III) and expressive (EVT) standardized vocabulary tests. They also responded to the fast mapping task in the comprehension (receptive) and production (expressive) modalities. It is well known that children may understand words before producing them during typical development. Yet Restrepo et al. (2006) observed just the opposite outcome. Their participants’ scores were higher on the expressive (EVT) than the receptive (PPVT–III) vocabulary tests. They suggested that the “open format of the EVT might be more conducive for AA children and children whose parents have lower education levels” (p. 25). In other words, the PPVT–III requires test takers to choose one of four picture options that match the examiner’s spoken vocabulary word. But the EVT allows test takers to use a number of words that match the meaning of targeted vocabulary items. Unlike Restrepo et al. (2006), the current study did not support an expressive or receptive mean score bias. A post hoc paired-samples t test revealed comparable mean scores for the PPVT–III and the EVT for the low-SES group, t(14) = 0.767, p = .455, d = 0.15 and for the middle-SES group, t(14) = 1.26, p = .227, d = 0.14. Small effect sizes were observed in each case.

Our findings may have differed from Restrepo et al. (2006) for two reasons. First, the characteristics of our participant sample were quite different. Our sample included younger children, who were typically developing and from two different SES backgrounds. Restrepo and colleagues focused on older preschool children who were developmentally at “high risk.” Second, differences in the parental level of education may be responsible for the discrepant findings between the two studies with respect to the receptive/expressive bias. In our sample, more than 70% of the parents had some technical training, college level education, or actual degrees compared with an estimated 56% of the AA parents in Restrepo et al. (2006).

Unlike the norm-referenced standardized tests, the fast mapping task did reveal a receptive/expressive bias in the current study. Toddlers from both economic groups demonstrated higher accuracy rates for the comprehension than the production of novel word targets. This observation suggested that the fast mapping task may tap into the same internal process that can explain why a child’s receptive vocabulary is typically larger than the expressive vocabulary and the comprehension of language occurs earlier than its production (Fenson et al., 1994). Recall that Carey (1978) proposed that during the initial fast mapping period, children are able to generate a lexical representation of a word’s meaning. Although children may be able to quickly generate a lexical representation after one or two exposures to a novel word, the strength of the representation may be more adequate for
recognition (comprehension) than production (expression) of the precise sequence of unfamiliar phonemes composing the novel words. If so, then even weak or partial representation would result in better comprehension than production performance measures.

**Research Implications**

Given the results of the current study, SES should continue to be examined in designing studies concerned with the vocabulary knowledge of AA children. While a growing number of studies have focused on AA children’s performance on norm-referenced vocabulary measures, there is comparatively less research on the fast mapping tasks. More research on fast mapping tasks is warranted given its potential clinical merit and what it might eventually tell us about the underlying processes involved in children’s word learning. Future research should replicate this study’s paradigm of multiple task comparisons on the same participant sample. The current study’s small participant sample may have affected the results of the fast mapping measure in particular, due to the large amount of individual variability observed in the accuracy rates. It is possible that a small sample size combined with large score variability led to the absent SES group difference observed as opposed to the processing nature of the fast mapping task, as argued. Therefore, future studies involving larger participant samples are needed to confirm the validity of our results.

In addition to increasing the number of children, future research needs to focus more broadly on the validity of the fast mapping task itself by paying particular attention to the type of task conditions that can mitigate or restrict the range of successful performances. The observed significant differences between the comprehension and production tasks for both SES groups suggest that fast mapping outcomes are susceptible to task conditions. Future research can determine whether the higher comprehension than production scores reflected an inherent modality bias or a presentation order effect, given that the comprehension task always followed the production task in the current study. Irrespective of such future outcomes, the very fact that a modality bias was observed at all suggests that performance on a processing task like fast mapping can be influenced by task conditions. The more processing demands made on performance, the greater is the possibility that prior experience could differentially influence performance outcomes. Thus, another question to raise is whether changes in the number and type of novel words taught would alter the reliability or validity of the fast mapping results.

Finally, future research should be designed to determine whether fast mapping tasks have advantages over more standard assessment measures for identifying early language delay among AA children. Poverty is a risk factor for language delay (Olswang, Rodriguez, & Timler, 1998). Comparing the fast mapping performance of typically developing AA toddlers from low-SES backgrounds to a group of AA toddlers with documented slow-developing language ought to contribute to the kind of information needed to clinically differentiate below average scoring children with typical and atypical development.

**Clinical Implications**

Fast mapping tasks may offer an alternative assessment procedure for those children, particularly low-SES, whose clinical status is difficult to determine. Such is the case when one cannot be sure that below average test scores result from the lack of relevant background exposure or experience due to low SES, a cultural difference, or a real clinical deficit. If the outcomes of the current study were confirmed by future research, then we would expect AA children with typically developing processing abilities to perform like other typically developing children on novel learning tasks, regardless of whether their low scores on experience-dependent measures were due to low SES or to cultural differences. Children with language delay or impairment would be expected to perform below average on fast mapping tasks and on other more conventional experience-dependent measures.

The current study suggests that even the conventional norm-referenced and language sample measures also may be useful tools for assessing AA children who are very young, and particularly those from middle-SES homes. This study clearly showed that the latter group’s performances were comparable to or exceeded those expected for typically developing children at the same ages on both norm-referenced and language sample measures. Even though our low-SES group used significantly fewer words than did the middle-SES group, its NDW mean ($M = 113$) was still comparable to the NDW mean ($M = 112$) reported for typically developing children at the same age in Watkins et al. (1995). The NDW mean for our middle-SES group was even higher than that reported in Watkins et al. (1995). Both SES groups in our study also had remarkably higher mean NDWs than those observed for late-talking toddlers, whose means ranged from 5.9 to 6.4 (Robertson & Ellis Weismer, 1999). Regardless of their lower scores relative to their middle-SES peers, toddlers in the low-SES group would not be viewed as atypically developing children based on the diversity of their vocabulary use in spontaneous speech. Similarly, their generally lower mean scores on the norm-referenced vocabulary tests relative to their middle-SES peers were not in the range of atypical performance either. These children typically scored within 1 SD of the normative group’s mean score of 100 with few exceptions, their mean scores averaging above 90 on both the receptive and expressive standardized tests. Such observations suggest that much can be learned about AA children’s vocabulary by using conventional measures at a very young age. It is unknown whether the same kind of results would apply at older ages, when children’s linguistic knowledge is more invested in their specific cultural experiences.

**Acknowledgments**

This project was funded by a minigrant through Project LASER. Project LASER (H326M000002) is a funded cooperative agreement of the Office of Special Education Programs, U.S. Department of Education. Development of the fast mapping task was supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC03731 awarded to the second author. The authors extend thanks to the parents, children, and day care providers in the East Knoxville, TN, area who participated in the study.
References


Hollingshead, A. B. (1975). Four factor index of social status. Unpublished manuscript, Yale University, New Haven, CT.


Olswang, L., Rodriguez, B., & Timler, G. (1998). Recommending intervention for toddlers with specific language learning difficulties: We may not have all the answers, but we know a lot. American Journal of Speech-Language Pathology, 7(1), 23–32.


Received September 22, 2005
Revision received April 30, 2006
Accepted July 5, 2007
DOI: 10.1044/1058-0360(2007/041)

Contact author: RaMonda Horton-Ikard, Department of Communication Disorders, Florida State University, 407 Regional Rehabilitation Center, Tallahassee, FL 32303. E-mail: rhorton2@fsu.edu.