

AUTHOR: Julie A. Washington; Holly K. Craig

TITLE: Performances of At-Risk, African American Preschoolers on the Peabody Picture Vocabulary Test—III

SOURCE: Language, Speech, and Hearing Services in Schools v30 no1 p75-82 Ja 1999

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited.

ABSTRACT

This study examined the performance of 59 at-risk, African American preschoolers on the Peabody Picture Vocabulary Test—III (PPVT—III, Dunn & Dunn, 1997). The subjects were considered at-risk based on low-income status and/or social status variables such as family density and teenage parents. A mean standard score of 91 and a standard deviation of 11 were achieved by these children. Although these scores are below those reported for the PPVT—III standardization sample, the performance spread resulted in a normal distribution of scores. Differences in performance based on gender and income were not apparent, but level of education of the primary caregiver significantly influenced performance. The findings indicate that unlike the Peabody Picture Vocabulary Test—Revised (PPVT—R, Dunn & Dunn, 1981) the PPVT—III is a culturally fair instrument that is appropriate for use with this population.

KEY WORDS: African American, assessment, vocabulary, at risk

The critical need for the identification and development of assessment instruments for use with children who present cultural and linguistic differences has been well-documented (Campbell, Dollaghan, Needleman, & Janosky, 1997; Terrell & Terrell, 1993; Vaughn-Cooke, 1986; Washington, 1996). The absence of non-biased language assessment instruments not only has a profound impact on our ability to identify children with language deficits, but, by implication, impedes attempts to provide appropriate intervention services for these children. In the case of African American children, the need for assessment instruments that are not biased culturally, and that do not penalize them for the use of African American English (AAE), is long-standing (Stockman, 1986; Terrell & Terrell, 1993; Vaughn-Cooke, 1986; Washington, 1996). In addition, African American children in the United States are disproportionately poor (U.S. Department of Commerce & U.S. Bureau of Census, 1997), which is one of the factors that can place children at risk. Poverty affects the use and acquisition of language skills in important ways (Fazio, Naremore, & Connell, 1996; Hart & Risley, 1995; Walker, Greenwood, Hart, & Carta, 1994) that may be evident in their performance on standardized tests.

Two major strategies have emerged to address the acute need for assessment procedures that are appropriate to the young African American child. One strategy has been to develop nonstandardized assessment taxonomies that are largely unaffected by AAE (Craig, 1996, Craig & Washington, 1994; Craig, Washington, & Thompson-Porter, 1998a, 1998b). For example, Craig and Washington (1994) presented a complex syntax scoring taxonomy and Craig et al. (1998a, 1998b) identified two comprehension tasks and average C-unit lengths as nonstandard assessment approaches that are suitable for assessing African American children. In addition, Campbell et al. (1997) identified nonword repetition as a culturally fair procedure for assessing African American children that does not penalize impoverished children for lack of world knowledge.

A second strategy undertaken to address the need for nonbiased assessment measures has been the modification of widely used standardized language instruments in an effort to reduce bias (Hemingway, Montague, & Bradley, 1981; Washington &

Craig, 1992; Wiener, Lewnau, & Erway, 1983). These modifications primarily have involved awarding credit for test items that are potentially impacted by dialect use or other cultural influences on responding. The former strategy of creating nonstandardized assessments is ongoing and has been successful and well-received, but the latter strategy involving the modification of existing instruments has been largely unsuccessful.

A notable attempt at modification that was unsuccessful was the second revision of the Peabody Picture Vocabulary Test (PPVT, Dunn, 1959). Although the Peabody Picture Vocabulary Test—Revised (PPVT—R, Dunn & Dunn, 1981) reflected a number of positive changes that were designed to reduce the cultural bias identified for the original PPVT (Dunn, 1959), it was determined to be inappropriate for use with most young children from different cultural and linguistic communities (Prewitt Diaz, 1988; Sharpley & Stone, 1985), including African American children (Washington & Craig, 1992).

Washington & Craig (1992) administered the PPVT—R to a sample of low-income, African American kindergartners and preschoolers and found that the test was racially or economically biased. Despite the application of a scoring adjustment for items that were missed by more than half of their subjects, they determined that there was not enough performance spread on the PPVT—R to provide useful information concerning the vocabulary skills of their young, African American subjects. A test that is appropriate for a given segment of the population should yield a range of performances from well above average to below average that closely approximates a standard normal distribution. Instead, Washington and Craig determined that the distribution of scores for their subjects was significantly skewed toward the low tail of the standard normal distribution (see Figure 1). Washington and Craig suggested that the PPVT—R test developers did not evaluate the scores of different minority groups separately, and consequently, performance differences that existed by race or socioeconomic status were not apparent.

Whereas the successful development of nonstandardized assessment protocols that can be used with African American children has been encouraging, the absence of standardized instruments for use with these children has been equally discouraging. Both clinicians and researchers report a need for valid standardized instruments that can be administered as a part of language screenings or assessment batteries (Tomblin et al., 1997), but the development of new instruments has not been forthcoming. The recent revision of well-established instruments that attempt to address the need raises the possibility that measures that were once considered inappropriate for use with African American children may be changed in ways that render them culturally fair. The purpose of this investigation was to explore the appropriateness of the Third Edition of the Peabody Picture Vocabulary Test (PPVT—III, Dunn & Dunn, 1997) for use with a population of at-risk, African American preschoolers.

The standardization procedures for the PPVT—III included more attention to possible differences in responding by race and gender (Williams & Wang, 1997). Specifically, a national tryout was conducted in which all items were administered to subjects representing three major racial/ethnic groups in this country: African American, Hispanic, and Native American. Based on the results of item analyses by group, 75 items were determined to be biased and were subsequently eliminated (Williams & Wang). In addition, based on the U.S. Census population estimates, gender and race representations were appropriate, and in some cases, over-representation was achieved. Average education level was chosen to represent socioeconomic status (SES), and all levels were reportedly well-represented (Williams & Wang).

Despite this attention to socioeconomic differences and race at the level of test construction, the technical manual for the PPVT—III still does not report performance

data for these potentially important influences on language use and development. Lack of attention to possible performance differences by race and SES during construction of an instrument raises critical questions concerning the applicability of the normal distribution to children of differing cultural, linguistic, and economic backgrounds (Fazio et al., 1996; Washington & Craig, 1992). It will be important to determine whether the procedures employed by the test developers to improve the reliability and validity of the PPVT—III have made it more appropriate for use with individuals representing varied race and income groups. This investigation examined the appropriateness of this instrument for use with urban African American children who were identified as being at risk for academic failure based on either income or other environmental concerns.

METHOD

SUBJECTS

The subjects were 59 African American boys ($n = 25$) and girls ($n = 34$) who ranged in age from 47 to 57 months (mean = 51 months). The subjects were identified from four state-sponsored, at-risk preschool classrooms in the Metropolitan Detroit area. There were a total of 78 4-year-old children enrolled in the preschool classrooms, 67 of whom were African American. Eight of the African American preschoolers did not show up for testing, despite repeated attempts to reschedule appointments.

A number of medical and social status variables made up the at-risk criteria that were used to select children for enrollment in the preschool. All of the subjects in this investigation were considered to be at risk by virtue of family incomes that were below the poverty line or social status variables such as family density, single parent households, and/or significant family histories (e.g., teenage parents). None of the subjects was considered to be at risk based on medical factors.

Fifty-five of the African American subjects were typically developing based on teacher and parent judgments, and none were enrolled in special education services of any type at the time of data collection, even though special education programs were available for children of this age. The Triangles subtest of the Kaufman Assessment Battery for Children (KABC, Kaufman & Kaufman, 1983), a cognitive test of nonverbal concept formation, was administered to each of these children and confirmed their normal developmental status (see Table 1). Four of the 59 African American children (2 boys and 2 girls) were on the school's special education caseload, and their development was considered atypical. Administration of the Triangles subtest confirmed that these four children were not performing within the normal range (see Table 1).

In addition, two language measures were available for each child from a language screening that was conducted at the beginning of the school year by a team of six certified speech-language pathologists. The measures included a language comprehension task eliciting responses to wh-questions and a language production task sampled during picture description. The Wh-Question Comprehension task involved presentation of two action pictures (picture numbers 33 and 35) from the Bracken Concept Development program (Bracken, 1986). Each picture was accompanied by 12 randomly ordered question forms that differed in level of complexity from simple object naming to comprehension of temporal and causal relationships (see Craig et al., 1998b for a complete description of this task). The picture description task involved the presentation of three pictures (picture numbers 11, 28, and 30), which were also from the Bracken Concept Development program (Bracken). The three longest C-units produced during picture description were identified for each child and the mean length of C-unit in words (MLCU-w) was calculated for screening purposes. Table 1 provides a summary of the performance of all subjects on these two measures. Examination of

responses on these two language measures established that all children were AAE speakers.

Family income was determined from school records and varied from poverty through middle-income levels as defined by the U.S. Census Bureau statistics for African American citizens (U.S. Department of Commerce & U.S. Bureau of Census, 1997).(FN1) This information was available for 51 of the typically developing subjects. In addition, the level of education of the children's primary caregiver was obtained from school records for these same 51 children. Most (97%) of the children's mothers were identified as the primary caregiver, but for two children, a maternal grandmother served as primary caregiver. Among all 51 caregivers, 14% had achieved less than a high school education, 71% had at least a high school diploma, and 18% had college degrees.

African American children made up more than 75% of the student body in the school district participating in this investigation. Each child passed a bilateral hearing screening that was administered by either a pediatrician or the school district prior to enrollment in the preschool.

DATA COLLECTION AND SCORING

Form IIIB of the PPVT—III (Dunn & Dunn, 1997) was randomly selected from the two forms available and administered to each child individually in a room that was free from distractions. It was administered according to published guidelines by one of a team of six certified speech-language pathologists, two of whom were African American and four of whom were Caucasian. A t-test for independent samples revealed nonsignificant variations in PPVT—III standard scores across subjects relative to race of examiner [$t(57) = .641, p > .05$]

Each subjects' responses were scored according to the established scoring criteria. Raw scores were converted to standard scores and percentiles. Interrater agreement was established by having an independent observer, a certified speech-language pathologist with past research experience, recalculate a randomly selected subset of the raw scores, standard scores, and percentile ranks of 10% of the subjects. Scoring agreement was 100% between raters.

RESULTS

The mean standard score achieved for the typically developing children was 91, with a standard deviation (SD) of 11. This score corresponds to approximately the 27th percentile when compared to the PPVT—III normative sample. The mean standard score for the small sample of atypical children was 78.0, with a SD of 15.2, which corresponds to the 7th percentile. It must be noted that although the mean score of the atypical children was more than 1 SD below the mean of the PPVT—III, the range of scores for these children was wide (55-96). All subsequent analyses excluded these four children, and results are reported only for the 55 typically developing children.

Both the mean standard score and SD for the 55 typically developing children were below the established standard score mean of 100 and SD of 15 for the PPVT—III, yet still within normal expectations for this instrument. Despite the difference in the mean and SD achieved for these subjects from the PPVT—III standardization sample, there was a wide performance spread in this at-risk population. Visual inspection of the data suggested that the distribution of scores approximated the normal curve. Application of the One Sample Kolmogorov-Smirnov test of normality with Lilliefors correction confirmed this impression ($K-S(55) = .077; p > .200$). Our earlier findings for the PPVT—R indicated that the results were significantly skewed when they were administered to a population of at-risk kindergartners and preschoolers (Washington & Craig, 1992). In contrast, the distribution of PPVT—III scores in this investigation

resulted in a nonsignificant difference from the standard normal distribution (see Figure 1).

The relationships among standard scores on the PPVT—III and the subjects' language and cognitive scores were examined. Table 2 reports the Pearson product moment correlations and the associated probability levels attained. Because more than one relationship was being examined, the experiment-wise alpha level was divided by three (the number of measures examined) and the more conservative alpha level of .016 was used. Using this more conservative standard, performance on the PPVT—III did not correlate at a statistically significant level with the children's cognitive scores ($r = .23$; $p > .016$). However, performance on the PPVT—III did correspond significantly to performance on the language production and comprehension measures. There was a low, positive, statistically significant relationship to average C-unit lengths ($r = .42$; $p = .002$) and a moderately strong, positive, statistically significant correlation between the PPVT—III standard score and performance on the Wh-Question Comprehension task ($r = .56$, $p = .000$).

The relationship between standard scores on the PPVT—III and selected social status variables was also examined. There were no interaction effects between factors ($p > .05$) and no main effects for either gender [$F(1, 50) = 2.13$; $p > .05$] or income [$F(2, 50) = .341$; $p > .05$]. However, there was a significant main effect for caregiver education [$F(2,50) = 4.35$; $p = .020$]. Table 3 includes the PPVT—III standard scores based on income level and caregiver education. A Tukey-HSD post hoc analysis revealed a significant difference in the mean PPVT—III standard scores of subjects whose primary caregiver had less than a high school education and those whose parents had completed high school and/or college ($p < .05$). (see Table 3).

DISCUSSION

The reported changes in construction of the PPVT—III (Dunn & Dunn, 1997) had a significant, positive impact on the appropriateness of this test for assessment of receptive vocabulary with this urban, at-risk sample of African American preschoolers. Washington and Craig (1992) cautioned against the use of the PPVT—R with at-risk preschoolers and kindergartners because it lacked performance spread. Most of their subjects (91%) scored significantly below the mean, and evidence of item bias was reported. In contrast, the performance spread achieved by the at-risk, African American children in this investigation was not statistically different from the standard normal distribution (see Figure 1). There was also no evidence of difficulty with specific items on the PPVT—III. These results suggest that the PPVT—III should be informative as part of a language assessment for characterizing receptive vocabulary skills of African American children.

The mean standard score and the SD around which the subject's scores distributed were lower than the mean and SD reported for the PPVT—III. Specifically, the mean standard score for the subjects in this investigation was 91, with a SD of 11. The established mean standard score and SD for the PPVT—III are 100 and 15, respectively. The performance of our African American subjects was indicative of performance at the 27th percentile compared to the PPVT—III mean.

The literature on at-risk children suggests that language is frequently affected, as evidenced by lower scores on standardized tests. Most of this literature focuses on children who are born in poverty (Bruck & Tucker, 1974; Fazio et al., 1996; Hess & Shipman, 1965; Walker et al., 1994). These studies cite the environmental stress that is frequently present in impoverished homes as a significant factor contributing to poor language performance for these children. The subjects in this investigation were not all impoverished. In fact, 35% of the sample fell in the third fifth of the census and could

be characterized as solidly middle income. However, the presence of significant, substantiated environmental stress (e.g., teenage pregnancy and family density) was a requirement for enrollment in the preschool classrooms from which these children were identified. Thus, these children were considered to be at risk despite family incomes that could be characterized as middle income.

Fazio et al. (1996) questioned the appropriateness of the means established for standardized instruments when applied to at-risk children. They found that their impoverished subjects consistently scored below the mean on language tests. The findings of this investigation provide further evidence that these means may not be directly applicable to at-risk children, regardless of family income status. Most language tests rely on prior experience and knowledge for attainment of "average" scores, and this is certainly true of vocabulary tests. Performance at the mean on the PPVT—III assumes that children of comparable ages, despite differences in experience and world knowledge, will perform at comparable levels. Considered together, the lower sample mean and the normal distribution of scores suggests that the assumption of comparable performance for at-risk children and the PPVT—III normative sample is not valid, but that the lower mean may more accurately characterize the vocabulary skills of these children at preschool age.

Examination of the gender, family income status, and level of education of the primary caregiver provided additional information concerning potential factors that might contribute to the differences in performance on the PPVT—III. Characterizing children as at risk provides only a general description that is not informative for describing performance-based differences. Gender (Craig & Evans, 1991; Ely, Berko-Gleason, & McCabe, 1996; Macaulay, 1978; McCarthy, 1930; Sheldon & Rohleder, 1996; Winitz, 1959), family income status, and caregiver education (Duncan, Brooks-Gunn, & Klebanov, 1994; McLoyd, 1990; Van Baar & DeGraff, 1994) have all been identified as important influences on children's language performance. No differences in performance by gender were apparent for the children in this investigation; the boys performed comparably to the girls (see Table 1). Most recent research finds few differences in performance by gender, and then primarily in discourse behaviors (Craig & Evans, 1991; Sheldon & Rohleder, 1996). Standard scores also did not differ significantly based on family income (see Table 3). In fact, the subjects' mean standard scores were quite comparable, with similar variability occurring between income groups.

Education level of the caregiver resulted in significant differences in performance, with the children whose caregivers had not completed high school achieving significantly lower scores than those children whose caregivers had completed either high school or college (see Table 3). These findings are consistent with those reported by Van Baar and DeGraff (1994) and Brooks-Gunn, Klebanov, and Duncan (1996). These investigators identified maternal education as the most significant "co-variable" influencing their African American subjects' language and academic achievement. Brooks-Gunn et al. (1996) reported that caregiver education was a highly significant predictor of poverty and performance among minority children in general and African American children in particular. The findings in this investigation confirmed the importance of caregiver education for distinguishing performance differences on the PPVT—III for at-risk, African American preschoolers.

The attention to test construction issues that were identified as shortcomings for the PPVT—R, including item analysis and the elimination of items based on performance differences by race, appears to have overcome many of the shortcomings that were reported for earlier versions of this instrument. However, the differences in performance based on caregiver education for the at-risk children in this investigation were unexpected based on the data presented in the PPVT—III technical manual (Williams

& Wang, 1997). The technical manual indicates that the normative sample consisted of children of different races, and that education level was used to index socioeconomic status. However, no performance data are reported relative to education level.

The results of this investigation suggest that differences in performance can be expected based on caregiver education and that the PPVT—III is sensitive enough to detect those differences. These outcomes suggest that future research might be informed by examining language performance differences of children from at-risk families based not only on income differences, but also on other, more informative factors such as level of education of the caregiver. In order to fully understand this population, it will also be important to identify additional co-variables that might positively or negatively impact performance.

A comparison of the subject's performance on the PPVT—III and additional language and cognitive scores revealed that there was no correlation between PPVT—III scores and performance on a nonverbal cognitive measure, but significant positive correlations with a language comprehension and a language production task. Nonverbal cognitive tests theoretically are designed to examine cognitive ability without the influence of language. The lack of significant relationship between scaled scores on the Triangles subtest of the KABC and the PPVT—III suggested that this subtest does in fact assess this discrete cognitive skill relatively independent of language functioning. Of the two language measures, the strongest relationship existed between the PPVT—III and the Wh-Question Comprehension measure. We would expect this because both instruments examine receptive language ability. The correlation between the language measures and the PPVT—III seemingly underscores the success of the PPVT—III in appropriately assessing the language abilities of the African American children in this investigation.

Further, these data for the PPVT—III and the Triangles subtest, when considered together with the Wh-Question task (Craig et al., 1998b), have the potential to offer clinicians a range of speech and language performances that are non-overlapping. These instruments are the beginning of a valid and culturally fair assessment for young, African American children.

ADDED MATERIAL

Julie A. Washington

Holly K. Craig

University of Michigan, Ann Arbor

Contact author: Julie A. Washington, PhD, University of Michigan, Communicative Disorders Clinic, 1111 E. Catherine St., Ann Arbor, MI 48109-2054. Email: julieaw@umich.edu

ACKNOWLEDGMENT

This investigation was supported by research grant #1 RO3 DC02884-01A1 from the National Institute on Deafness and Other Communication disorders, National Institutes of Health. The authors thank Carol Connor and Kathleen Welch for their significant contributions to various aspects of the project. Special thanks also to Alexander Bailey, Leon Crumb, Patricia McCarthy, and the early childhood teachers and students in the Oak Park, Michigan STAR program.

Table 1. KABC scaled scores (SS), total scores on the Wh-Question Comprehension task and average C-unit lengths (MLCU-w) on the Picture Description task for 59 subjects.

| Typically developing (n = 55) | | Atypical (n = 4) | |
|----------------------------------|----|---------------------|----|
| M | SD | M | SD |

| | | | | |
|--------------------------------|--------------|------|--------------|-------|
| SS | 10.72 (FN*) | 1.71 | 6.50 | .58 |
| Wh-Question (FN _a) | 50.13 (FN**) | 7.44 | 23.00 (FN**) | 13.78 |
| MLCU-w | 5.71 (FN***) | 1.79 | 3.33 (FN***) | 2.61 |

FOOTNOTES

* $t(57) = 4.88$; $p < .001$;

** $t(57) = 6.62$; $p < .001$;

*** $t(57) = 2.49$; $p < .05$.

a Total points possible on the Wh-Question Comprehension task was 72.

Table 2. Correlations among PPVT-III standard scores, Wh-Question Comprehension (Wh-q), scaled scores of the triangles subtest of the KABC, and average C-unit lengths (MLCU-w)

| | PPVT-III | Wh-q | Triangles | MLCU-w |
|----------|----------|------------|-----------|------------|
| PPVT-III | -- | .56 | .23 | .42 |
| P | | .000 (FN*) | .112 | .002 (FN*) |

FOOTNOTE

* $p \leq .016$.

Table 3. PPVT-III standard score means and standard deviations (SD) by income group and level of caregiver education.

| | Mean | SD |
|---------------------------|-------------|------|
| Total Sample (n = 55) | 91.0 | 11.0 |
| Income groups | | |
| Third fifth | 91.00 | 11.6 |
| Second fifth | 89.92 | 8.0 |
| Bottom fifth | 89.62 | 13.0 |
| Caregiver education (FN*) | | |
| College graduate | 94.0 | 12.3 |
| High school graduate | 93.2 | 8.8 |
| < high school | 77.3 (FN**) | 10.7 |

FOOTNOTES

* $[F(2,50) = 4.35$; $p < .05]$

** $p < .05$.

Figure 1. The standard normal distribution, the distribution of PPVT-R scores for 105 low-income African American children, and the distribution of PPVT-III scores for the current at-risk, African American subjects.

FOOTNOTE

1 The Statistical Abstract of the United States (U.S. Department of Commerce & U.S. Bureau of Census, 1997) reports incomes based on the aggregate income received by each fifth and the top 5 percent of families nationwide, by race and Hispanic origin. Based on the data provided for "Black" heads of household, the incomes of the families whose children participated in this investigation were representative of the lowest, second, and third fifths. Accordingly, these families presented incomes ranging from less than \$8,055 to an upper limit of \$50,000. More specifically, 41% of the families had incomes in the lowest fifth, 24% in the second fifth, and 35% in the third fifth.

REFERENCES

Bracken, B. A. (1986). Bracken Concept Development Program. San Antonio, TX: Psychological Corp.

Brooks-Gunn, J., Klebanov, P. K., & Duncan, G. J. (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. *Child Development*, 67, 396-408.

- Bruck, M., & Tucker, G. (1974). Social class differences in the acquisition of school language. *Merrill Palmer Quarterly*, 20, 205-220.
- Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech and Hearing Research*, 40(3), 519-525.
- Craig, H. (1996). The challenges of conducting language research with African American children. In A. G. Kamhi, K. E. Pollock, & J. L. Harris (Eds.), *Communication development and disorders in African American children* (pp. 1-17). Baltimore, MD: Brookes Publishing.
- Craig, H., & Evans, J. (1991). Gender and children's turn-taking. *Journal of Speech and Hearing Research*, 34, 866-878.
- Craig, H., & Washington, J. (1994). The complex syntax skills of poor, urban, African American preschoolers at school entry. *Language, Speech, and Hearing Services in Schools*, 25(3), 181-190.
- Craig, H., Washington, J., & Thompson-Porter, C. (1998a). Average C-unit lengths in the discourse of African American children from low-income, urban homes. *Journal of Speech, Language, and Hearing Research*, 41(2), 433-444.
- Craig, H., Washington, J., & Thompson-Porter, C. (1998b). Performances of young African American children on two comprehension tasks. *Journal of Speech, Language, and Hearing Research*, 41(2), 445-457.
- Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Child Development*, 65, 296-318.
- Dunn, L. (1959). *Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service.
- Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Test—Revised*. Circle Pines, MN: American Guidance Service.
- Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.
- Ely, R., Berko-Gleason, J., & McCabe, A. (1996). Why didn't you talk to your mommy honey? Gender differences in talk about past talk. *Research on Language in Social Interaction*, 29, 7-25.
- Fazio, B., Naremore, R. C., & Connell, P. J. (1996). Tracking children from poverty at risk for specific language impairment: A 3-year longitudinal study. *Journal of Speech and Hearing Research*, 39, 611-624.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes Publishing.
- Hemingway, B. L., Montague, J. C., & Bradley, R. H. (1981). Preliminary data on revision of a sentence repetition test for language screening with Black first-grade children. *Language, Speech, and Hearing Services in Schools*, 12, 145-152.
- Hess, R., & Shipman, V. (1965). Early experience and the socialization of cognitive modes in children. *Child Development*, 36, 869-886.
- Kaufman, A., & Kaufman, N. (1983). *Kaufman Assessment Battery for Children: Interpretive manual*. Circle Pines, MN: American Guidance Service.
- Macaulay, R. (1978). The myth of female superiority in language. *Journal of Child Language*, 5, 353-363.
- McCarthy, D. (1930). The language development of the preschool child. In Y. Brackbill (Ed.), *Infancy and early childhood* (pp. 3-14). New York: The Free Press.
- McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socioeconomic development. *Child Development*, 61, 311-346.
- Prewitt Diaz, J. O. (1988). Assessment of Puerto Rican children in bilingual education programs in the United States: A critique of Lloyd M. Dunn's monograph.

Special issue: Achievement testing: Science vs. ideology. *Hispanic Journal of Behavioral Sciences*, 10(3), 237-252.

Sharpley, C. F., & Stone, J. M. (1985). An exploratory investigation to detect cross-cultural differences on the PPVT-R. *Psychology in Schools*, 22(4), 383-386.

Sheldon, A., & Rohleder, L. (1996). Sharing the same word, telling different stories: Gender differences in co-constructed pretend narratives. In D. I. Slobin, J. Gerhardt, A. Kyratzis, & J. Guo (Eds.), *Social interaction, social context, and language* (pp. 613-632). Mahwah, NJ: Lawrence Erlbaum Associates.

Stockman, I. J. (1986). Language acquisition in culturally diverse populations: The Black child as a case study. In O. L. Taylor (Ed.), *Nature of communications disorders in culturally and linguistically diverse populations* (pp. 117-155). San Diego, CA: College-Hill Press.

Terrell, S. L. & Terrell, F. (1993). African American cultures. In D. E. Battle (Ed.), *Communication disorders in multicultural populations* (pp. 3-37). Boston, MA: Andover Medical Publishers.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 40, 1245-1260.

U.S. Department of Commerce, & U.S. Bureau of the Census. (1997). *Statistical abstract of the United States* (116th ed.). Washington, DC: Authors.

Van Baar, A., & DeGraff, B. M. T. (1994). Cognitive development at pre-school age of infants of drug dependent mothers. *Developmental Medicine & Child Neurology*, 36(12), 1063-1075.

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, MA: College-Hill Press.

Walker, D., Greenwood, C. R., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. Special issue: Children and poverty. *Child Development*, 65(2), 606-621.

Washington, J. (1996). Issues in assessing language skills in African American children. In A. G. Kamhi, K. E. Pollock, & J. L. Harris (Eds.), *Communication development and disorders in African American children: Research assessment and intervention* (pp. 35-54). Baltimore, MD: Brookes Publishing.

Washington, J., & Craig, H. (1992). Performances of low-income, African American preschool and kindergarten children on the Peabody Picture Vocabulary Test-Revised. *Language, Speec and Hearing Services in Schools*, 23, 329-333.

Wiener, F. D., Lewnau, L. E., & Erway, E. (1983). Measurin language competency in speakers of Black American Englis *Journal of Speech and Hearing Disorders*, 48, 76-84.

Williams, K., & Wang, Y. (1997). *Technical manual: Peabody Picture Vocabulary Test-III*. Circle Pines, MN: American Guidance Service.

Winitz, H. (1959). Language skills of male and female kindergtten children. *Journal of Speech and Hearing Research*, 2, 377-386.

Received February 10, 1998

Accepted July 14, 1998