Reliability and Validity of the Black English Sentence Scoring System

Yvette D. Hyter
Western Michigan University

8-1984

Follow this and additional works at: http://scholarworks.wmich.edu/masters_theses

Part of the Speech Pathology and Audiology Commons

Recommended Citation
http://scholarworks.wmich.edu/masters_theses/1490

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master's Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
RELIABILITY AND VALIDITY
OF THE BLACK ENGLISH SENTENCE SCORING SYSTEM

by

Yvette D. Hyter

A Thesis
Submitted to The
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Arts
Department of Speech Pathology and Audiology

Western Michigan University
Kalamazoo, Michigan
August 1984
The purpose of this study was to assess the concurrent validity and interscorer reliability of a specific diagnostic tool of language assessment which is designed to be nonbiased for dialectally different children. Black English Sentence Scoring (BESS; Nelson, 1983), which is based on Lee's (1974) Developmental Sentence Scoring (DSS), is a system for analyzing spontaneous language samples that gives credit for the normally developing features of Black English dialect. The sample consisted of 17 children between three and seven years who were residing in Michigan in communities where Black English was spoken frequently. All were identified as having language disorders by their speech-language pathologists. The investigator administered two criterion tests, Screening Kit of Language Development (SKOLD) and Structured Photographic Expressive Language Test (SPELT) to each child and gathered spontaneous language samples, which were transcribed and analyzed using BESS/DSS. The investigator and faculty advisor each scored the language samples as a means of establishing interscorer reliability. BESS was found to correlate with SPELT and SKOLD and interscorer reliability was found to be high.
ACKNOWLEDGEMENTS

So many were instrumental in the instigation and completion of this study. I wish to express my gratitude to those listed below.

First and foremost, I give thanks to God for blessing me with the courage and strength to continue through the most difficult stages of this study.

Dr. Nickola W. Nelson, my faculty advisor, was most inspirational and never tiring. She provided encouragement and guidance during the course of this study.

I would like to thank my thesis committee, Mrs. Shirley Sparks, Dr. Michael Clark, and Dr. Michael Payne, for their helpful suggestions and guidance. I would like to thank Jonelle Robinson for assisting in gathering and transcribing the language samples as well as the many speech-language pathologists who assisted in identifying children on their caseloads.

Without the cooperation of the Kalamazoo and Detroit, Michigan, public school administrations, I would not have been able to carry out this study. I would like to thank principals of the schools involved for allowing me to use their facilities. I would also like to thank the children and their parents who were willing to participate.

Finally, I would like to thank my family for their undying support, encouragement and enthusiasm.

Yvette D. Hyter
INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages ______
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy √
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages √
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine ______
10. Computer printout pages with indistinct print ______
11. Page(s) ___________ lacking when material received, and not available from school or author.
12. Page(s) ___________ seem to be missing in numbering only as text follows.
13. Two pages numbered ___________. Text follows.
14. Curling and wrinkled pages _____
15. Other __________________________________________

University Microfilms International
# Table of Contents

Acknowledgements .............................................. ii

List of Tables ................................................ v

Chapter

I. INTRODUCTION ........................................ 1

   Statement of Purpose ................................. 3

   Experimental Hypotheses ............................ 4

II. REVIEW OF THE LITERATURE ............................ 5

   Nonbiased Assessment ................................ 6

   Development of Black English ....................... 8

   Features of Black English ........................... 9

   Formal Language Tests ............................... 15

   Language Sampling as a method of Nonbiased Assessment.. 15

III. METHOD .............................................. 18

   Subjects ............................................ 18

   Assessment ......................................... 19

   Black English Sentence Scoring in Conjunction with Developmental Sentence Scoring ......................... 20

   Procedure ........................................ 25

   Variables Affecting Study ........................... 27

IV. RESULTS ............................................ 29

V. DISCUSSION AND CONCLUSIONS ......................... 40

   Conclusions Regarding Experimental Hypotheses ........ 41

   Limitations of the BESS ............................ 50

   Summary ............................................. 51
LIST OF TABLES

1. Description of Tests ................................................................. 24
2. Mean and Standard Deviation Values of Subjects ............. 30
3. Rank Difference Correlation: BESS and SPELT ............... 33
4. Rank Difference Correlation: BESS and SKOLD ............... 35
5. Pearson Product Moment Correlation: BESS and SPELT ..... 36
6. Pearson Product Moment Correlation: BESS and SKOLD ..... 37
7. Pearson Product Moment Correlation: SPELT and SKOLD ..... 38

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Chapter I

Introduction

Standardized norm referenced testing is often used by speech-language pathologists as a basis for determining language functioning of all individuals, including those of diverse cultural and language backgrounds. This information is then used to establish the degree of speech and language disorder. However, when children of diverse cultures* are evaluated, they may appear to be language disordered because they have language skills that do not match those measured by most tests (Arnold and Reed, 1976; Dale, 1972; Dillard, 1972; Hubbell, 1981; Smitherman, 1977; Taylor and Payne 1983; Terrell and Terrell, 1983; Wolfram and Fasold, 1974). This cultural bias in testing leads some children to be inappropriately considered to be language disordered and to be wrongly placed in special education programs.

When testing the language of children who are culturally different, special considerations must be made for the fact that differences may exist in any of the systems of language content, form or use (Bloom and Lahey, 1978). Simply translating a test into the form and vocabulary of another dialect does not ensure that the test reflects the same topics and functions of language that those different children use (Adler and Birdsong, 1983; Vaughn-Cooke, 1983).

*A number of terms used in reporting this study may require clarification of meaning. A list of such terms and their definitions as they are used here are provided in Appendix A.
Traditional concepts of test reliability and validity must be expanded when testing culturally different children who may be language disordered. The tests must reflect the kinds of skills a person might be expected to have if developing normally, and since children of different cultures might be expected to have different experiences, tests may need to measure different content (Adler and Birdsong, 1983; Vaughn-Cooke, 1983). Tests must be standardized on the population for which they will be used (Adler and Birdsong, 1983; Weiner and Hoock, 1973). They also must be shown to differentiate children who are developing language normally from those who are not (Dale, 1972; Hubbell, 1981).

Several methods are available for testing language skills. One such method is gathering spontaneous language samples. Gathering language samples and analyzing them is an accepted procedure for conducting evaluations of language disordered children (Bloom and Lahey, 1978; Lee, 1974; Leeman, 1981; Taylor and Payne 1983; Vaughn-Cooke, 1983). For culturally different children, spontaneous language sampling offers a method more free of bias than use of standardized tests, in that children are more free to select their own ideas and to use language in a variety of forms. Unfortunately, until recently, there has been an absence of reliable and valid tools that accurately assess the language of culturally different children (Taylor and Payne, 1983; Terrell and Terrell, 1983; Vaughn-Cooke, 1983).

A Black English Sentence Scoring (BESS) system, which is designed to be used in conjunction with Developmental Sentence Scoring (DSS; Lee, 1974) has been developed (Nelson, 1983). However, its
ability to differentiate language disorder from normally developing
language of children who speak Black English has not yet been demon-
strated. The purpose of this study is to investigate further the
concurrent validity and interscorer reliability of this instrument.

Statement of Purpose

When children of diverse cultures who do not have language dis-
orders are evaluated with standardized norm referenced tests, they
may appear to be language disordered because they have language skills
that are different from the Standard English norm. This cultural bias
in testing leads some children to be inappropriately considered to be
speech-language disordered. The lack of appropriate test instruments
also makes it difficult to diagnose dialectally different children
who do have speech-language disorders.

The purpose of this study was to assess the validity of a speci-
fic diagnostic tool of language assessment which is designed to be
nonbiased for dialectally different children. Black English Sentence
Scoring (BESS; Nelson, 1976, 1983) is a system for analyzing sponta-
neous language samples that gives credit for normally developing
features of Black English dialect. Standardization data have been
gathered for normal children between the ages of three and seven years
(Nelson, 1983), but the validity of the BESS as an assessment tool has
yet to be established. This study was designed to measure the con-
current validity and interscorer reliability of the BESS system.
Experimental Hypotheses

The general hypothesis to be tested was that the Black English Sentence Scoring system in conjunction with the Developmental Sentence Scoring system of language analysis would yield scores that would support other diagnostic findings. The following hypotheses were established:

1. That the children who had been identified, prior to the study by an Individual Education Planning Committee (IEPC) or comprehensive assessment, as language disordered would also (a) score below suggested cut off scores on the criterion measures, Screening Kit of Language Development (SKOLD; Bliss and Allen, 1983) and Structured Photographic Expressive Language Test (SPELT; Werner and Kreseck, 1983) and would (b) be found to be language disordered using BESS;

2. That there would be a significant rank-difference correlation coefficient between scores on the BESS and the criterion measures (SKOLD and SPELT);

3. That significant positive correlation (at the .05 level of confidence) would be found between (a) BESS and SKOLD scores, and (b) BESS and SPELT scores, using Pearson product moment correlation coefficient; and

4. That the percent of interscorer agreement between two examiners on four randomly selected BESS transcripts would be high (90 percent or greater).
CHAPTER II

REVIEW OF THE LITERATURE

For sometime, a controversy existed as to whether Black English represented a language difference or a language disorder. Although Black English was once thought to be indicative of illiteracy, a position identified as the "deficit theory", it is now recognized widely as a legitimate language system. The "difference theory" in which Black English is viewed as representing normal variation in language from community to community (Wolfram, 1979) is a widely supported one (Baratz, 1969; Committee on CCC Language State, 1975; Dillard, 1972; Labov, 1966, 1969; Shuy, 1969; Wolfram, 1970; Wolfram and Fasold, 1974).

Attempts to apply the difference theory clinically, however, have led to questions about which methods should be used in distinguishing different language from disordered language (Terrell and Terrell, 1983). An additional problem has been the lack of methods for adequately identifying language disorders in dialectally different individuals (Vaughn-Cooke, 1983). Unfortunately, the development of nonbiased or culturally fair evaluation measures has been limited (Taylor and Payne, 1983; Terrell and Terrell, 1983; Vaughn-Cooke, 1983).
Diagnosis of communication disorders in general requires that children be classified, in part, according to their performance on standardized tests. Thus, the "accuracy of the classification depends on the accuracy of the validity of the instrument" (Seymour and Seymour, 1979, p.459).

For the dialectally different child, validity may need to be determined separately. For example, if a test of grammatical Standard English structures were administered to Standard English speaking children, their scores might reflect their abilities. However, if the same test were administered to Black English speaking children without modification, it might be invalid for measuring their language abilities (Seymour and Seymour, 1979).

It is important to have valid measures to assess the language skills of all children, whether they are language different or disordered, to ensure that all will receive appropriate education and special education services if, and only if, they are needed. One needs to consider the relevance of the content of the test to the child's culture, and the test must be administered only to the children of the culture from which the norms for the test were obtained (Adler and Birdsong, 1983).

In 1979, the controversy surrounding these problems was debated in federal court (Bountress, 1980). The "Ann Arbor (King) Decision," as it has come to be known, dealt with a number of issues concerning the differences of Black English and the performance of speakers of
Black English on standardized tests and learning as a whole. The courts ruled that teachers need to become aware of the possibility of children with differing dialects in their classrooms and they need to attend to the special needs of those children who have differing dialects (Bountress, 1982; Monteith, 1980; Robinson, 1981; Smitherman, 1977; Yellin, 1981).

In the Education for All Handicapped Children Act of 1975 (Ballard, 1975), it is mandated that evaluation materials used for the purpose of assessment and placement of children must be selected so as not to be racially or culturally discriminating. Terrell and Terrell (1983) have described varied responses to this mandate. Some speech-language pathologists have apparently responded to this mandate by operating on the assumption that all minority group children are normal dialect speakers, and professionals in this category have often chosen not to evaluate minority children for possible language disorders unless severe problems exist. Other speech-language pathologists do evaluate children selected through referral or screening processes, but they use standardized tests and then operate on the assumption that any communicative behaviors that differ from those expected are characteristic of the child's different language style and should be viewed as normal. Terrell and Terrell (1983) call both versions of this process "overcompensation." Still other speech-language pathologists continue to practice "undercompensation" and assume that any child's language which differs from the expected responses on standardized tests is disordered (Terrell and Terrell, 1983).
In attempts to control for both overcompensation and undercompensation, it has been suggested that speech-language pathologists must (1) become familiar with the cultural style of the groups to which individuals being evaluated belong, in that they must (2) maintain an appropriate attitude toward that group and (3) that they must become critically aware of test standardization data (Adler and Birdsong, 1983; Terrell and Terrell, 1983).

Development of Black English

There is little research on language acquisition related to dialect differences. However, there is a "common stage of language development that all children go through regardless of the adult dialect they will speak" (Wolfram, 1979, p. 6). Language development is the "gradual and progressive integration" of three language components: content (what people say); form (the mode or manner of expression); and use (why people speak and the selection of forms in varied situations) (Bloom and Lahey, 1978, p. 10). The differences between Black English and Standard English, are most prominent in the component of language form (phonology, morphology and grammar) (Hood, 1973; Vaughn-Cooke, 1983; Ward, 1971), and that is the component of Black English dialect that has heretofore received the most study and comment.

When a child is in the beginning of his/her language developmental period, language differences are not found (Wolfram, 1979). Yet, as the child observes and monitors the language of his/her models, language differences appear between cultures (Hubbell, 1981; Wolfram, 1979). A child learns and uses the language his/her environment reinforces.
Features of Black English

Black English is recognized as a legitimate form of language (Dillard, 1972; Wolfram and Fasold, 1974) and contains its own particular linguistic system (Thomas, 1972). The following discussion of Black English features will be divided into two parts: (1) phonology and (2) grammar. This information is taken from Fasold, 1970; Reveron, 1981; Smitherman, 1977; Williams, 1972; Wolfram, 1970; and Wolfram and Fasold, 1974.

Phonology

Phonology refers to "the units of sound" (Bloom and Lahey, 1978, p. 15). Phonology is discussed here because it has morphological significance. Six phonological features of Black English are discussed; (a) consonant cluster reduction, (b) the th sounds, (c) r-lessness and l-lessness, (d) devoicing of final stops b, d, g, (e) nasalization, and (f) vowel glides.

(a) consonant cluster reduction

Final consonant cluster reduction is found when, (1) both ending sounds are voiced or voiceless (e.g., tes/"test," build/"build"); (2) the following word begins with a vowel (e.g., col'egg/"cold egg"); and (3) the -ed suffix produces voiced or a voiceless cluster (this affects ed as a past tense marker, e.g., She play outside/"She played

In each of the examples in this section of the text, Black English versions are followed by Standard English of the same forms in quotation marks.
outside"). Consonant clusters are also altered during plural forma-
tion. For words ending in -sp, -st, and -sk, the -es is used instead
of -s (e.g., deses/"desks," ghoses/"ghosts"). The consonant cluster
is present when the suffix begins with a vowel (e.g., tester, coldest).

(b) -th sounds

Two -th sounds will be discussed: The voiceless -th as in "no-
ting" and the voiced -th as in "bathe". There is a -f and -t sub-
stitution of the voiceless -th in the middle and final position of
words (e.g., nufin/"nothing," baf/"bath," sumtin/"something," wit/"with"); a -d for the voiced -th in the initial and middle position
of words (e.g., day/"they," bruder/"brother"); and a -v for the voiced
-th in the final position of words (e.g., bave/"bathe").

(c) r-lessness and 1-lessness

r may be absent in the intervocalic position (e.g., Ma'y/"Mary," sto'y/"story"). Dropping of -r at the end of words affects pronoun
usage (e.g., you/"your," they/"their"). Therefore, these forms would
be produced as "It is they book," or "It is you book". -r may be ab-
sent when it follows a consonant (e.g., p'otect/"protect") or when
the following vowel is either an -o or -u (e.g., th'ow/"throw"). The
-1 becomes uh after a vowel (e.g., steauh/"steal"). The absence of
-1 may affect contracted forms (e.g., Tomorrow I bring it/"Tomorrow
I'll bring it"); -1 is absent when it precedes w,r,1 (e.g., aweek/
"all week"); and deletion of -1 does not occur in the intervocalic
position (e.g., bullet, alive).

(d) devoicing of final stops b,d,g

At the end of syllables or words, voiced stops, b,d,g, are pro-
duced like the voiceless counterparts, p,t,k. The vowels that occur before a voiced consonant, are produced with longer duration in Black English. It is this feature that discriminates the Standard English pronunciation of "but" and the Black English pronunciation of "but" for "bud".

(e) nasalization

The use of in for ing is present in both Standard English and Black English (e.g., singin'/'singing"). However, it is used more frequently in Black English. The use of a nasalized vowel at the end of words instead of a nasal consonant is common in Black English. The nasal consonant is dropped producing words such as mā/'man" and bū/"bun".

(f) vowel glides

Vowel glides such as ei and ai are pronounced without the glide (e.g., boah/'boy," toah/'toy"). The absence of the glide frequently occurs more often when the glide is followed by a voiced consonant (absent in side, and time, rather than kite and flight). This, among others, is a feature that is also found in some parts of the South in non-Black English speakers.

Grammar

The term grammar refers to the units of sound, units of meaning and the way they interact with each other; i.e., the rules of language (Bloom and Lahey, 1978). Ten grammatical features of Black English will be discussed: (a) past tense forms, (b) perfective constructions, (c) the third person singular present tense marker, (d) future,
(e) invariant be, (f) absence of forms to be, (g) negation and multiple
negation, (h) possessive, (i) plural, and (j) pronominal apposition.

(a) past tense forms

Because of the consonant reduction rule, the -ed suffix, which
marks the past tense, past participle, and derived adjective, is not
pronounced (e.g., finish/"finished"). However, when the suffix -ed
is added to a word ending in -t or -d, it is pronounced (e.g., batted,
waded).

(b) perfective constructions

The present perfective forms in Black English are, I walked (I
have walked), whereas in Standard English the form is, I've walked,
(I have walked). The past perfect form in Black English is the same
as in Standard English (e.g., I had walked). However, Standard English
speakers also use the contracted form (e.g., I'd walked). A completive
form (e.g., I done walked) and a remote time form (e.g., I been walked)
are present only in Black English. The completive form denotes an
action started and completed at a specific time in the past (e.g., I
done tried). Remote time construction indicates that the speaker thinks
of action having taken place in the distant past. This construction
is used with "been", (e.g., I been had it).

(c) the third person singular present tense marker

The -s suffix in Black English is absent (e.g., he walk/"he walks")
when making the third person singular in the present tense. The ab­
sence of the -s suffix affects the use of the verb do as an auxiliary
in negative constructions (e.g., He don't go/"He doesn't go"). Third
person forms has and does are absent (e.g., He have a bike/"He has
a bike" and He always do silly things/"He always does silly things").
Some Black English speakers use hypercorrection forms of the -s suffix at times. The speaker uses the -s suffix to mark the first, second, and third person forms both singular and plural forms (e.g., I walks, you walks, they walks).

(d) future

Gonna is a future indicator. When gonna is used, is and are are frequently omitted (e.g., He gonna go). There are three reductions of gonna; I'mana go, I'mon go, and I'ma go. Will is also used as an indicator of the future. The contracted form of will (I'll) may be eliminated in Black English, especially if the following word begins with a labial consonant (e.g., I bring it tomorrow/"I'll bring it tomorrow").

(e) invariant be

There are five finite different forms of to be; am, is, are, was and were. In Black English, the form be is used as the main verb (e.g., I be here this evening). The contraction of will is often absent before the form be indicating future tense (e.g., He be here pretty soon). The contracted form of would may be omitted by the final elimination rule of final consonants described above (e.g., If you gave him a present, he be happy). Another form of invariant be is when there is no tense specification and seems to describe an object or an event distributed intermittently in time (e.g., He be mad when somebody leave him home).

(f) absence of forms to be

When is and are are expected in Standard English, they may be absent in Black English. Yet, is and are are present in exposed clauses (e.g.,
I know he is).

(g) negation and multiple negation

Ain't corresponds to Standard English' use of didn't, isn't and aren't in some varieties of Black English. Multiple negation occurs when one negative is expressed in more than one place in a sentence (e.g., He doesn't know nothing). The word any does not appear in a negative sentence. If a sentence has an idenfinite noun phrase containing a negative marker (e.g., nobody, nothing) before a verb, the negativized form of the verbal auxiliary (e.g., can't, wasn't) may be placed at the beginning of the sentence producing a sentence like "Can't nobody do it".

(h) possessive

Black English speakers may indicate possession by the order of words, whereas in Standard English, -'s possessive is used (e.g., The boy hat/"The boy's hat"). The -'s possessive may also be inappropriately used with personal names, representing a form of hypercorrection (e.g., John's Jackson cat or John's Jackson's car).

(i) plural

Most people who speak Black English use plural markers, but such markers (-s or -es) are occasionally absent in Black English grammar. In Standard English some nouns are pluralized by vowel change (e.g., one foot, two feet) or with no change at all (e.g., one deer, two deer), In Black English those nouns often are used with the regular -s suffix (e.g., two foots, two deers).

(j) pronominal apposition

Pronominal apposition is the construction in which a pronoun
occurs immediately after the noun subject of the sentence. The nomi-
native form of the pronoun is used (e.g., My brother, he bigger than
you). Wolfram (1969) states that this feature is used by all speakers
of Standard and Black English. However, Lee (1974) treated it as an
immaturity among Standard English speakers in designing her DSS sy-
stem. Speakers of Black English are given credit for this feature on
the BESS system (Nelson, 1983).

Formal Language Tests

A major problem in using criterion measures to establish the
concurrent validity of a test is in identifying appropriate criterion
procedures. One can use clinical judgement to identify children with
language disorders, but the opportunity to measure statistical corre-
lations is then lost. In validating a measure of proficiency that
includes Black English, criterion measures must be identified that
are themselves appropriate for speakers of Black English (Vaughn-
Cooke, 1983). Two such tests that have been identified as appropriate
concurrent measures for the present study are: Screening Kit of
Language Development (SKOLD; Bliss and Allen, 1983) and Structured
Photographic Expressive Language Test (SPELT; Werner and Kresheck,
1983). The SKOLD and SPELT will be discussed in greater detail in
the section on Evaluation Instruments.

Language Sampling as a Method of Nonbiased Assessment

Gathering language samples as a method of language assessment is
a widely used and accepted procedure (Bloom and Lahey, 1978; Lee, 1974;
Taylor and Payne, 1983; Tyack and Gottsleben, 1974; Vaughn-Cooke, 1983). However, one must analyze results with caution. Vaughn-Cooke (1983) states that language samples alone cannot provide critical information needed to make a diagnosis regarding the "normalcy" of a child's language skill and the results obtained from a language sample should be carefully interpreted. Two language sampling techniques, Developmental Sentence Scoring and Black English Sentence Scoring, will be discussed here.

Developmental Sentence Scoring (DSS; Lee, 1974) is a method for scoring and evaluating a child's use of Standard English grammatical rules from a tape recorded sample of the child's spontaneous speech to an adult (Lee, 1974). Because this method of language analysis is based only on Standard English grammatical rules, the "occurrence of Black English features is penalized" (Nelson, 1983, p. 4).

Black English Sentence Scoring (BESS; Nelson, 1983) is a system of language analysis. It was devised to "assist in differential diagnosis" (Nelson, 1983, p. 4) of language disorder from language differences in children who have learned to speak or who are learning to speak Black English. In using BESS, a spontaneous language sample is gathered from a child and is analyzed using the same grammatical categories as DSS (Nelson, 1983). The categories are: indefinite pronoun, personal pronoun, primary verb, secondary verb, negatives, conjunctions, interrogative reversals, and WH-questions (Lee, 1974). The scoring system of BESS differs in that credit is given for Black English features even when they differ from Lee's (1974) guidelines for scoring Standard English. During the development of BESS, test-
retest reliability and construct validity were demonstrated. However, concurrent validity and interscorer reliability of this instrument have yet to be established.
CHAPTER III

METHOD

Subjects

The subjects for the investigation were identified with the assistance of speech-language pathologists within the Kalamazoo and Detroit area who were willing to participate in the study. Those professionals assisted in identifying children on their caseloads who met the criteria for the study outlined below. One child was identified through the Kalamazoo Headstart Program, four were enrolled in Kindergarten in the Kalamazoo Public School System and 12 were placed in a special Preprimary program for language impaired children within the Detroit Public School System. Of the 17 subjects, 14 were male and three were female.

All subjects were between the ages of 3;8 and 6;11 and were identified by interview and teacher report as children who were using, or learning to use, Black English and who resided in homes or communities where Black English occurred frequently. All were identified as having language disorders through Individual Education Planning Committee (IEPC) process (except for the Headstart child, who was identified using comprehensive assessment process), and their speech-language pathologists also described them as having specific language disorders involving problems in the form of language (word endings, phoneme usage, word order). The speech-language pathologists used
various measures to identify the language disorders and were sensitive to cultural and linguistic differences associated with Black English usage. The language disorders did not result from any of the primary etiological conditions, mental retardation, sensorineural hearing loss, autism, and/or neuromuscular disease. All subjects were verbal enough to produce 50 sentences within one hour.

All subjects were intelligible and cooperative. The parents of each subject were informed of the nature of the study in a letter and through a telephone conversation with the investigator. Their written permission for their child to participate was obtained prior to the initiation of the evaluation measures. Children and parents were informed of their rights to withdraw from the study without penalty at any time during its progress and were informed that no individually identifiable information would be revealed to anyone other than the primary investigator, the supervisor of the study and the assisting speech-language pathologist. The Western Michigan University Human Subjects Review Board agreed with the above terms before the initiation of this study.

Assessment

Evaluation Instruments

Three instruments were used in this study to assess the language skills of the subjects; Black English Sentence Scoring (BESS) in conjunction with Developmental Sentence Scoring (DSS); Screening Kit of Language Development (SKOLD); and Structured Photographic Expressive
Language Test (SPELT). In addition to these instruments, a good quality tape recorder with remote microphone was used to record the language samples. Stimulus materials to aid in the language elicitation component of the evaluation included stories, pictures, and toys.

Black English Sentence Scoring in Conjunction with Developmental Sentence Scoring

When using Black English Sentence Scoring (BESS) in conjunction with Developmental Sentence Scoring (DSS), a spontaneous language sample of at least 50 sentences was gathered using toys, pictures, and stories. The sample was recorded on audiotape for later transcription. Once the language sample had been gathered and transcribed, it was analyzed using Lee's (1974) DSS system and Nelson's (1983) BESS system.

Only complete consecutive utterances were analyzed. According to Lee (1974), a complete sentence consists of a noun and a verb occurring in a subject-predicate relationship. However, in BESS a sentence is complete if it also is constructed with a "zero" (deleted) copula that occurs with a predicate noun or a predicate adjective (Nelson, 1983). "She my mom", and "That boy naughty", are examples.

The BESS system uses the same eight categories as the DSS system, yet allows credit for Black English features. The eight categories include: indefinite pronoun, personal pronoun, primary verb, secondary verb, negatives, conjunctions, interrogative reversals, and WH-questions (See Appendix B for Scoring System). Each sentence is analyzed twice, once using DSS methods and another using BESS methods (See Appendix C for example of analysis).
**Screening Kit of Language Development**

The Screening Kit of Language Development (SKOLD; Bliss and Allen, 1983), is used as an early identification measure of children with language disorders between the ages of 2;6 and 4;0 years. It includes normative data for children who speak Black English, as well as for those who speak Standard English. This test consists of six subtests, S30, S37, S43 (Standard English) B30, B37, B43 (Black English) to screen children 30-36 months, 37-42 months, and 43-48 months, respectively. The SKOLD tests comprehension and production of words and sentences through vocabulary comprehension, story completion, sentence comprehension, paired sentences with pictures, individual sentences with pictures, individual sentences without pictures, and comprehension of commands.

The SKOLD takes approximately 15 minutes to administer. One is not to skip any subtests and must give each in order. The test is discontinued when the minimum score for passing (at the top of the score sheet) has been reached by the subject. If a child does not pass the test with a minimum passing score, then he/she should be referred to a speech-language pathologist for further testing. For this study, the entire test was administered so that a total score could be obtained for each subject.

Interobserver, interexaminer and intraexaminer reliability were reported for the SKOLD (Bliss and Allen, 1983). All measures of reliability were obtained with a correlation of 0.85 or more. Concurrent validity was also obtained. The criterion measures for the SKOLD

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
were the **Sequenced Inventory of Communication Development (SICD)** and clinician judgement (Bliss and Allen, 1983).

**Structured Photographic Expressive Language Test**

The **Structured Photographic Expressive Language Test (SPELT; Werner and Kresheck, 1983)** is divided into two parts. The SPELT-P is for preschoolers 3;0 to 4;11, and the SPELT-II is for children 4;0 to 8;11.

The SPELT-P and SPELT-II are designed as screening instruments that allow identification of children who may have difficulty in expressing morphological and syntactic features. Both tests include responses considered acceptable for speakers of Black English. The SPELT uses colorful photographs as visual stimuli. The auditory stimuli consist of statements or questions that the examiner says when the pictures are presented. The response is then given by the child.

The morphological aspects measured include content words (nouns, verbs), function words (prepositions, pronouns, auxiliary verbs), and inflected endings (noun and verb suffixes). The syntactic aspect measured is word order (the correct sequence of words in a sentence). The SPELT-II allows analysis of a child's ability to use several common grammatical forms (singular and plural nouns; subject; object; possessive and reflexive pronouns; main verbs; copulas; auxiliary verbs; secondary verbs; prepositions; conjunctions; negatives and contractions). The results of these tests are intended to indicate the need for further diagnosis of language if necessary.

Test-retest, internal consistency, and interscorer reliability
were reported for the SPELT-II (Werner and Krescheck, 1983). Correlation coefficients of .85 - .88 were obtained for test-retest reliability for ages four through six (see Table 1 for summary of tests).
<table>
<thead>
<tr>
<th>TEST</th>
<th>Age Range</th>
<th>Time to Administer</th>
<th>Time to Score</th>
<th>Items Evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>BESS</td>
<td>3;0-6;11</td>
<td>30 min. to one hour</td>
<td>13 1/2 min.</td>
<td>Form of Language - Indefinite pro. - Personal pronouns - Primary verb - Secondary verb - Negatives - Conjunctions - Interrogative reversals - WH-questions</td>
</tr>
<tr>
<td>DSS</td>
<td>3;0-6;11</td>
<td>30 min. to one hour</td>
<td>13 1/2 min.</td>
<td>Same as Above</td>
</tr>
<tr>
<td>SKOLD</td>
<td>2;6-4;0</td>
<td>15 min. to 25 min.</td>
<td>5 min. or as presenting items</td>
<td>Comprehension and Production - Vocabulary - Story completion - Sentence Compre. - Paired sentences - Indiv. sentences - Comprehension of commands</td>
</tr>
<tr>
<td>SPELT-P</td>
<td>3;0-4;11</td>
<td>10 min.</td>
<td>5 to 7 min. or as presenting items</td>
<td>Morphology and Syntax - Content words - Function words - Inflected endings - Word order</td>
</tr>
<tr>
<td>SPELT-II</td>
<td>4;0-8;11</td>
<td>15 min.</td>
<td>5 to 7 min. or as presenting items</td>
<td>Morphology and Syntax - Grammatical forms</td>
</tr>
</tbody>
</table>
Low correlation coefficients (.30) were obtained for ages seven and eight. A reliability value of .70 was obtained for internal consistency reliability and an interscorer correlation of .99. The criterion measures that were used to report test validity were the Test of Language Development (TOLD) and Developmental Sentence Scoring (DSS) (Werner and Kresheck, 1983). There was 100 percent agreement between the SPELT and the TOLD and 82 percent agreement between SPELT and the DSS. The same reliability measures were obtained for SPELT-P as for SPELT-II. However, test-retest reliability resulted in a correlation coefficient of .94; internal consistency reliability resulted in a reliability value of .70; and interscorer reliability resulted in 100 percent agreement in scoring. The Preschool Language Scale and DSS were used as criterion measures for testing the validity of the SPELT-P. A correlation of .76 indicated similarities in performance on these tests (Werner and Kresheck, 1983).

Procedure

Test Administration

The investigator, with the assistance of another individual (Jonelle Robinson (J.R.), Detroit Public Schools) administered the SKOLD first, followed by the SPELT and then gathered a language sample from each of the subjects. The SKOLD and SPELT were administered prior to the language sample in an effort to put the child at ease with the interaction format. The SKOLD was administered prior to the SPELT because it is a longer test. The entire evaluation procedure took
approximately one to two hours. Testing of some subjects was spread over two sessions (within a two week period) because of the time required to reduce the likelihood of the child becoming fatigued.

Reliability

The investigator (Y.H.) and J.R. each scored all of the tests (SKOLD, SPELT, BESS) that they had administered. However, as a measure of reliability between scorers, the faculty advisor (N.N.), re-scored four randomly selected language samples from the 17 subjects. Prior to this procedure, a training phase was instituted, in which 25 sentences from three language samples of the normative study (Nelson, 1983) were scored by Y.H. and then compared with the results gathered in 1983. The interscorer reliability measurements were based on the number of times Y.H. and N.N. agreed on the exact score for individual items on the four randomly selected experimental samples. This was expressed as a percentage of the total number of different scores.

Validity

Validity refers to the ability of a test to measure what it purports to measure. Several types of validity exist: concurrent (the obtained results correlate positively with scores on a previously validated test), content (the items of the test appear to the examiner to be a reasonable representative sample of the skills the test is supposed to measure, also called face validity), predictive (the test provides a prediction of some future outcome), and construct (reveals the characteristics the examiner has that causes the subject to respond
in a particular way) (Amos, Brown and Mink, 1965; Anastasi, 1961; Leeman, 1981). The purpose of this study was to analyze the concurrent validity of the BESS system, using SKOLD and SPELT as criterion measures. Construct validity has been supported by the increments on scores that accompanied increments of age in the normative study (Nelson, 1983).

Concurrent validity refers to comparison of performance on one test with other widely used measures (Anastasi, 1961; Leeman, 1981). In this case the scores of the BESS in conjunction with DSS were compared with scores obtained with other measures (SKOLD and SPELT) which have been adapted for Black English speakers.

Variables Affecting Study

A number of variables that might possibly affect results of this study were considered in designing it. Two of the major variables are socioeconomic status of the subject and experimenter bias.

Socioeconomic Status of Subject

The subjects of this study were primarily of lower socioeconomic status. Studies have found that Black individuals of lower socioeconomic status are more likely to speak Black English (Wolfram, 1969; Wolfram and Fasold, 1974). However, the studies are questionable due to the use of a small number of subjects (Baratz, 1970).

The current study included some subjects of lower socioeconomic status but subjects were not necessarily limited to lower socioeconomic status. An area was provided on the permission slip (See Appendix D)
for parents to report educational and employment status. These were rated according to a social class scale (Warner, 1960) (See Appendix E), however there was no attempt to control the variable of socioeco-
onomic status.

**Experimenter Bias**

The investigator of this study was aware that all of the children in the study had already been identified as language disordered. This knowledge may have influenced the investigator's objectivity in testing these individuals. However, it was the relative correlation of the scores on the multiple measures that was the sole measure of validity. Interscorer reliability procedures were used to reduce the likelihood that experimenter bias would influence the results. After Y.H. transcried and scored the language samples, N.N., who was not aware of the identities or ages of the children, but did know that the children were language disordered, scored the language samples again.
CHAPTER IV

RESULTS

Validity of a test is impossible to establish unless it is also determined to be reliable (Silverman, 1977). Split-half and test-retest reliability of the Black English Sentence Scoring (BESS; Nelson, 1976, 1983) system were measured during the normative study (Nelson, 1983). Concurrent validity and interscorer reliability were measured in this study. The subjects' scores on the BESS and criterion measures were arranged by rank-difference correlation as one means of measuring test validity. The consistency of the results was also tested with the Pearson product moment correlation coefficient and the \( t \)-test for Pearson \( r \). Measures for establishing interscorer reliability were taken and they will be described in this chapter, as well.

The raw scores obtained on the BESS/DSS (Developmental Sentence Scoring; Lee, 1974), and the two criterion measures (Screening Kit of Language Development (SKOLD) and Structured Photographic Expressive Language Test (SPELT), and a description of the subjects are reported in Appendix F. The range of scores on the SKOLD was 10 to 35 points with a mean of 28 points. The range on the SPELT was 10 to 39 points with a mean of 25 points. The range on the BESS was 4.61 to 9.28 with a mean of 7 points. A table of mean DSS, BESS, SKOLD, and SPELT values and standard deviations at one year intervals has been provided (See Table 2), and comparative scores with normative samples are included, where available.
**TABLE 2**

Mean and Standard Deviation Values of Subjects

<table>
<thead>
<tr>
<th>Age</th>
<th>N</th>
<th>Mean DSS</th>
<th>S.D.</th>
<th>Mean BESS</th>
<th>S.D.</th>
<th>Lee's (1974) Mean DSS Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>3;0 to 4;0</td>
<td>1</td>
<td>4.42</td>
<td>--</td>
<td>5.20</td>
<td>--</td>
<td>6.64</td>
</tr>
<tr>
<td>4;0 to 5;0</td>
<td>2</td>
<td>3.5</td>
<td>--</td>
<td>5.24</td>
<td>--</td>
<td>8.04</td>
</tr>
<tr>
<td>5;0 to 6;0</td>
<td>8</td>
<td>4.8</td>
<td>1.44</td>
<td>6.98</td>
<td>1.44</td>
<td>9.19</td>
</tr>
<tr>
<td>6;0 to 7;0</td>
<td>6</td>
<td>5.2</td>
<td>.88</td>
<td>6.95</td>
<td>.94</td>
<td>10.94</td>
</tr>
</tbody>
</table>

Mean and Standard Deviation Values of Subjects from Normative Study (Nelson, 1983)

<table>
<thead>
<tr>
<th>Age</th>
<th>N</th>
<th>Mean DSS</th>
<th>S.D.</th>
<th>Mean BESS</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3;0 to 4;0</td>
<td>16</td>
<td>5.68</td>
<td>1.03</td>
<td>7.57</td>
<td>1.06</td>
</tr>
<tr>
<td>4;0 to 5;0</td>
<td>16</td>
<td>7.48</td>
<td>1.76</td>
<td>9.09</td>
<td>1.46</td>
</tr>
<tr>
<td>5;0 to 6;0</td>
<td>16</td>
<td>8.58</td>
<td>1.99</td>
<td>10.40</td>
<td>2.13</td>
</tr>
<tr>
<td>6;0 to 7;0</td>
<td>16</td>
<td>9.30</td>
<td>2.10</td>
<td>11.13</td>
<td>1.88</td>
</tr>
</tbody>
</table>
BESS: Identifying Children as Language Disordered

For identification of children as language disordered in this study, "failure" was defined as a subject's score: (1) falling below the cut off scores on the SPELT and SKOLD, and/or (2) falling one standard deviation or more below the mean of BESS. Informal observation revealed three of the 17 subjects failed the SKOLD, one subject passed the SPELT, three passed BESS and one passed DSS. All subjects were expected to be clearly identified as language disordered (failed) on the DSS because Black English features are not credited during this procedure, and DSS may thus identify even normally developing Black English speaking children as language disordered. Because the BESS allows points for Black English features, the BESS scores were expected to be higher than the DSS scores but still within the range of one standard deviation or more below the mean. The subjects were expected to pass the SKOLD because all but one of the subjects were older than the upper age limit of this screening instrument. In fact, during computation of the SKOLD results, the pass/fail criteria were not used because the ages for the cut-off scores were too low for the subjects. The numerical value was therefore used instead. It was thought that this was justified since none of the children earned the maximum possible score.

The three children who failed the SKOLD (although two of them were above the ceiling age) also were identified as language disordered on the SPELT and the BESS. Fourteen of the subjects passed the SKOLD and all 14 were found to be language disordered on the SPELT. Twelve
of them were found to be language disordered on BESS. One subject passed the SPELT and was not identified as language disordered on the SKOLD; yet was identified as language disordered on the BESS. However, this subject was at the cut off score for passing the SPELT. The BESS found all but three of the children to be language disordered. All three of these children failed the SPELT and all passed the SKOLD (which might be expected, since they were above its ceiling age, it will be recalled). These subjects were found to miss items involving prepositions, present progressive, past copula, past irregular on the SPELT, yet used these features appropriately on the BESS in their spontaneous language samples.

Correlation Coefficients: Rank-Difference and Pearson r

In effort to further investigate the validity of the BESS a rank-difference correlation coefficient method, which is similar to Pearson product moment correlation (Smith, 1970), was used. A perfect correlation between two measures (BESS and SKOLD; BESS and SPELT) would be equal to 1.00 and a perfect inverse correlation would be equal to -1.00. If the result was .00, there would not be any relationship between measures.

When using the rank-difference correlation method, which is "best suited to studies involving a small number of cases, say 15-30" (Smith, 1970, p. 155), a high coefficient of .997 was revealed, which indicates a significant rank-difference correlation between the BESS and the SPELT (See Table 3). A rank-difference correlation coefficient of .677 was obtained for the BESS and SKOLD. This figure also
### TABLE 3

Rank Difference Correlation: BESS and SPELT

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>SPELT</th>
<th>BESS</th>
<th>Rsp</th>
<th>Rb</th>
<th>(D) ((Rsp - Rb))</th>
<th>(D^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>5.20</td>
<td>17</td>
<td>16</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>7.40</td>
<td>7.5</td>
<td>5</td>
<td>2.5</td>
<td>6.25</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>9.28</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>5.86</td>
<td>13</td>
<td>12</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
<td>8.26</td>
<td>7.5</td>
<td>3</td>
<td>4.5</td>
<td>20.25</td>
</tr>
<tr>
<td>6</td>
<td>39</td>
<td>9.14</td>
<td>11</td>
<td>2</td>
<td>9</td>
<td>81</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>6.58</td>
<td>10</td>
<td>8</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>38</td>
<td>7.66</td>
<td>2</td>
<td>4</td>
<td>-2</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>30</td>
<td>6.48</td>
<td>6</td>
<td>10</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>27</td>
<td>6.10</td>
<td>9</td>
<td>11</td>
<td>-2</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>18</td>
<td>4.61</td>
<td>12</td>
<td>17</td>
<td>-5</td>
<td>25</td>
</tr>
<tr>
<td>12</td>
<td>32</td>
<td>6.61</td>
<td>4</td>
<td>9</td>
<td>-5</td>
<td>25</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>6.64</td>
<td>11</td>
<td>7</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>14</td>
<td>34</td>
<td>7.22</td>
<td>3</td>
<td>6</td>
<td>-3</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>5.70</td>
<td>15</td>
<td>13</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>14</td>
<td>5.18</td>
<td>15</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>14</td>
<td>5.28</td>
<td>15</td>
<td>14</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

\(D^2 = \frac{1}{233.5}\)

\[r' = 1 - \frac{6(233.5)}{17(289-1)} = 1 - \frac{14.01}{4896} = 1 - .003 = .997\]

\[r' = 1 - \frac{6D^2}{n(n^2 - 1)}\]
indicates significant rank-difference correlation between the BESS and SKOLD (See Table 4), but one that is less strong.

Additional correlation coefficients were obtained through the use of the Pearson product moment correlation method (a more powerful test than the rank-difference correlation coefficient method). A t-test for the Pearson r was also used to find the level of significance of the results. When comparing the BESS with the SPELT using the Pearson product moment correlation coefficient, a coefficient (r value) of .81 with a t value of 5.85 was obtained. Inasmuch as a perfect correlation is 1.00 and perfect inverse correlation is -1.00, this method revealed a high correlation between the BESS and SPELT. The t value of 5.85 was greater than the tabulated probability value of 2.131 (degree of freedom was n-2=15) at the .05 level of confidence. This finding further indicates significant correlations between the BESS and SPELT (See Table 5).

When comparing the BESS and SKOLD using the Pearson product moment correlation coefficient method, an r value of .66 was obtained. This value is also statistically significant, but is not as consistent. The t value of 3.40 was significant at the .05 level of confidence (degree of freedom was n-2=15) (See Table 6).

Because of the difference in correlation coefficients between the BESS and SPELT and the BESS and SKOLD, the SPELT was compared to the SKOLD using the Pearson product moment correlation coefficient method. An r value of .929 and a significant t value (25.88) was obtained at the .05 level of confidence (See Table 7).
## TABLE 4

**Rank Difference Correlation: BESS and SKOLD**

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>SKOLD</th>
<th>BESS</th>
<th>Rsk</th>
<th>Rb</th>
<th>D</th>
<th>D^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>5.20</td>
<td>17</td>
<td>16</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>7.40</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>9.28</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>5.86</td>
<td>15</td>
<td>12</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>8.26</td>
<td>11.5</td>
<td>3</td>
<td>8.5</td>
<td>72.25</td>
</tr>
<tr>
<td>6</td>
<td>34</td>
<td>9.14</td>
<td>3.5</td>
<td>2</td>
<td>1.5</td>
<td>2.25</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>6.58</td>
<td>11.5</td>
<td>8</td>
<td>3.5</td>
<td>12.25</td>
</tr>
<tr>
<td>8</td>
<td>35</td>
<td>7.66</td>
<td>1.5</td>
<td>4</td>
<td>2.5</td>
<td>6.25</td>
</tr>
<tr>
<td>9</td>
<td>34</td>
<td>6.48</td>
<td>3.5</td>
<td>10</td>
<td>6.5</td>
<td>42.25</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>6.10</td>
<td>6</td>
<td>11</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>4.61</td>
<td>13</td>
<td>17</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>27</td>
<td>6.66</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>29</td>
<td>6.64</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>35</td>
<td>7.22</td>
<td>1.5</td>
<td>6</td>
<td>4.5</td>
<td>20.25</td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>5.70</td>
<td>14</td>
<td>13</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>26</td>
<td>5.18</td>
<td>10</td>
<td>15</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>5.28</td>
<td>16</td>
<td>14</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

\[
r' = 1 - \frac{6(D^2)}{n(n^2-1)} = 1 - \frac{6(263.5)}{17(289-1)} = 1 - \frac{1581}{4896} = 1 - .323 = .677
\]

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
TABLE 5

Pearson Product Moment Correlation: BESS and SPELT

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>SPELT Scores</th>
<th>BESS Scores</th>
<th>X 1</th>
<th>Y 2</th>
<th>X²</th>
<th>Y²</th>
<th>X²Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>5.70</td>
<td>-14.6</td>
<td>-1.46</td>
<td>213.16</td>
<td>2.13</td>
<td>21.32</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>7.40</td>
<td>4.4</td>
<td>.74</td>
<td>19.36</td>
<td>.55</td>
<td>3.26</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>9.28</td>
<td>6.4</td>
<td>2.62</td>
<td>40.96</td>
<td>6.76</td>
<td>16.77</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>5.86</td>
<td>8.6</td>
<td>.80</td>
<td>73.36</td>
<td>.64</td>
<td>6.88</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
<td>8.26</td>
<td>4.4</td>
<td>1.00</td>
<td>19.36</td>
<td>2.56</td>
<td>7.04</td>
</tr>
<tr>
<td>6</td>
<td>39</td>
<td>9.14</td>
<td>14.4</td>
<td>2.48</td>
<td>207.36</td>
<td>6.15</td>
<td>35.71</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>6.58</td>
<td>1.6</td>
<td>.08</td>
<td>2.56</td>
<td>.01</td>
<td>1.13</td>
</tr>
<tr>
<td>8</td>
<td>38</td>
<td>7.66</td>
<td>13.4</td>
<td>1.00</td>
<td>179.56</td>
<td>1.00</td>
<td>13.4</td>
</tr>
<tr>
<td>9</td>
<td>30</td>
<td>6.48</td>
<td>5.4</td>
<td>.18</td>
<td>29.16</td>
<td>.03</td>
<td>- .97</td>
</tr>
<tr>
<td>10</td>
<td>27</td>
<td>6.10</td>
<td>2.4</td>
<td>.26</td>
<td>5.76</td>
<td>.31</td>
<td>-1.31</td>
</tr>
<tr>
<td>11</td>
<td>18</td>
<td>4.61</td>
<td>6.6</td>
<td>-2.05</td>
<td>43.56</td>
<td>4.20</td>
<td>13.53</td>
</tr>
<tr>
<td>12</td>
<td>32</td>
<td>6.61</td>
<td>7.6</td>
<td>-.05</td>
<td>54.76</td>
<td>.00</td>
<td>- .37</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>6.64</td>
<td>4.6</td>
<td>-.02</td>
<td>21.16</td>
<td>.00</td>
<td>.09</td>
</tr>
<tr>
<td>14</td>
<td>34</td>
<td>7.22</td>
<td>9.4</td>
<td>.56</td>
<td>88.36</td>
<td>.31</td>
<td>5.26</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>5.70</td>
<td>10.6</td>
<td>-.96</td>
<td>122.26</td>
<td>.92</td>
<td>10.18</td>
</tr>
<tr>
<td>16</td>
<td>14</td>
<td>5.18</td>
<td>10.6</td>
<td>-.48</td>
<td>122.26</td>
<td>2.19</td>
<td>15.69</td>
</tr>
<tr>
<td>17</td>
<td>14</td>
<td>5.28</td>
<td>10.6</td>
<td>-.38</td>
<td>122.26</td>
<td>1.90</td>
<td>14.63</td>
</tr>
</tbody>
</table>

\[
\frac{\bar{X}=24.6}{\bar{X}=6.66} \quad \frac{\bar{X}=.20}{\bar{Y}=.02} \quad \frac{X^2=1336.12}{Y^2=29.76} \quad 163.89 = 2.68
\]

\[x^2 = 161.21 \quad r = \frac{161.21}{199.31} \approx 0.81 \]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Subject Number</th>
<th>SPELT Scores</th>
<th>BESS Scores</th>
<th>x</th>
<th>y</th>
<th>x^2</th>
<th>y^2</th>
<th>*xy</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>5.20</td>
<td>-16.1</td>
<td>-1.46</td>
<td>259.21</td>
<td>2.13</td>
<td>23.51</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>7.40</td>
<td>3.9</td>
<td>.74</td>
<td>15.21</td>
<td>.55</td>
<td>4.64</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>9.28</td>
<td>3.9</td>
<td>2.62</td>
<td>15.21</td>
<td>6.86</td>
<td>10.22</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>5.86</td>
<td>- 9.1</td>
<td>-.50</td>
<td>82.81</td>
<td>.64</td>
<td>7.28</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>8.26</td>
<td>- 1.1</td>
<td>1.60</td>
<td>1.21</td>
<td>2.56</td>
<td>-1.76</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>34</td>
<td>9.14</td>
<td>7.9</td>
<td>2.48</td>
<td>82.41</td>
<td>6.15</td>
<td>19.59</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>6.50</td>
<td>- 1.1</td>
<td>-.08</td>
<td>1.21</td>
<td>.01</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>35</td>
<td>7.66</td>
<td>8.9</td>
<td>1.00</td>
<td>79.21</td>
<td>1.00</td>
<td>8.3</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>34</td>
<td>6.48</td>
<td>7.9</td>
<td>-.18</td>
<td>82.41</td>
<td>.03</td>
<td>-1.42</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>6.10</td>
<td>1.9</td>
<td>-.56</td>
<td>15.21</td>
<td>.31</td>
<td>-2.12</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>4.61</td>
<td>5.1</td>
<td>2.05</td>
<td>26.01</td>
<td>4.20</td>
<td>10.46</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>27</td>
<td>6.61</td>
<td>.9</td>
<td>-.05</td>
<td>.81</td>
<td>.00</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>29</td>
<td>6.64</td>
<td>2.9</td>
<td>-.02</td>
<td>8.41</td>
<td>.00</td>
<td>.06</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>35</td>
<td>7.22</td>
<td>8.9</td>
<td>.56</td>
<td>79.21</td>
<td>.31</td>
<td>4.98</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>5.70</td>
<td>- 7.1</td>
<td>-.96</td>
<td>50.41</td>
<td>.92</td>
<td>6.32</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>26</td>
<td>5.18</td>
<td>- .1</td>
<td>-1.48</td>
<td>.01</td>
<td>2.19</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>5.28</td>
<td>-10.1</td>
<td>-1.38</td>
<td>102.01</td>
<td>1.90</td>
<td>13.74</td>
<td></td>
</tr>
<tr>
<td>Σx=26.1</td>
<td>Σx^2=6.66</td>
<td>Σy=19.98</td>
<td>Σxy=860.97</td>
<td>Σy^2=29.76</td>
<td>Σr^2=110.58</td>
<td>r=105.11</td>
<td>105.11</td>
<td></td>
</tr>
</tbody>
</table>

\[
x = 105.11 \\
y = 105.11 \\
r = 105.11 \\
\sqrt{\frac{\sum x^2 \cdot \sum y}{\sum x^2 \cdot \sum y}} = .66
\]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
### TABLE 7

Pearson Product Moment Correlation: SPELT and SKOLD

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>SKOLD</th>
<th>SPELT</th>
<th>X</th>
<th>Y</th>
<th>X^2</th>
<th>Y^2</th>
<th>XY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>10</td>
<td>-16</td>
<td>-15</td>
<td>256</td>
<td>225</td>
<td>240</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>29</td>
<td>4</td>
<td>4</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>31</td>
<td>4</td>
<td>6</td>
<td>16</td>
<td>36</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>16</td>
<td>-9</td>
<td>-9</td>
<td>81</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>29</td>
<td>-1</td>
<td>4</td>
<td>1</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>34</td>
<td>39</td>
<td>14</td>
<td>14</td>
<td>64</td>
<td>196</td>
<td>112</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>23</td>
<td>-1</td>
<td>-2</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>35</td>
<td>38</td>
<td>9</td>
<td>13</td>
<td>81</td>
<td>169</td>
<td>117</td>
</tr>
<tr>
<td>9</td>
<td>34</td>
<td>30</td>
<td>5</td>
<td>8</td>
<td>64</td>
<td>25</td>
<td>40</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>27</td>
<td>4</td>
<td>2</td>
<td>16</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>18</td>
<td>-5</td>
<td>-7</td>
<td>25</td>
<td>49</td>
<td>33</td>
</tr>
<tr>
<td>12</td>
<td>27</td>
<td>32</td>
<td>-1</td>
<td>7</td>
<td>1</td>
<td>49</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>20</td>
<td>-3</td>
<td>-5</td>
<td>9</td>
<td>25</td>
<td>-15</td>
</tr>
<tr>
<td>14</td>
<td>35</td>
<td>34</td>
<td>9</td>
<td>9</td>
<td>81</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>14</td>
<td>-7</td>
<td>-11</td>
<td>49</td>
<td>121</td>
<td>77</td>
</tr>
<tr>
<td>16</td>
<td>26</td>
<td>14</td>
<td>0</td>
<td>-11</td>
<td>0</td>
<td>121</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>14</td>
<td>-10</td>
<td>-11</td>
<td>160</td>
<td>196</td>
<td>117</td>
</tr>
</tbody>
</table>

\[ r = \frac{\text{sum of } XY}{\text{sum of } X \times \text{sum of } Y} = \frac{931}{1061 \times 1166} = 0.931 \]

\[ \sqrt{\frac{\text{sum of } X^2 \times \text{sum of } Y^2 - \text{sum of } X \times \text{sum of } Y^2}{\text{sum of } X^2 \times \text{sum of } Y^2}} = \sqrt{\frac{1001.9611}{1003926}} = 0.929 \]
Interscorer Reliability

In an effort to establish interscorer reliability, four language samples were randomly selected from the 17 subjects. Each was scored first by the primary investigator (Y.H.) and then independently by the faculty advisor (N.N.). The first 25 sentences from each transcript were used during the procedure. With each sentence there was an average of six possible scores which provided a total of 560 possible scores to be matched between two examiners.

The number of scores that were agreed upon between Y.H. and N.N. were added together and a percentage was taken from this total based on the 560 possible points. This method revealed 90 percent agreement between the two speech-language pathologists. This finding provides preliminary indication of acceptable interscorer reliability in scoring the BESS when a short training period of three to four hours has preceded the scoring.
CHAPTER V

DISCUSSION AND CONCLUSIONS

The purpose of this study was to assess the concurrent validity and interscorer reliability of the Black English Sentence Scoring (BESS) system. This system is a diagnostic tool of language assessment designed to be nonbiased for dialectally different children for whom Black English features are expected based on their cultural and sociolinguistic status. It is also used to analyze spontaneous language samples giving credit for normally developing features of Black English dialect.

When gathering subjects for this study, the investigator found few children on the caseloads of practicing speech-language pathologists in either Headstart or Public School settings that met the criteria (see p. 18 of text). In seeking subjects for this study, few children between 3;0 and 4;0 years of age were identified as language disordered. The children within this age range who were identified as language disordered, exhibited severe language disorders and were unable to produce 50 sentences within 45 minutes to an hour. This suggests that the process that Terrell and Terrell (1983) refer to as "overcompensation" may be operating. This is the process in which speech-language pathologists assume that all minority children are normal dialect speakers, or in which speech-language pathologists assume, after referral or screening processes using standardized tests, that although communicative behaviors...
differ from those expected, these behaviors should be viewed as normal.

Conclusions Regarding Experimental Hypotheses

The results and the conclusions that can be drawn for each of the four hypotheses researched during this study will be discussed here.

Hypothesis One

Hypothesis One stated that the children who had been identified before the study by an Individual Education Planning Committee (IEPC) or comprehensive assessment, as language disordered, would (a) score below suggested cut off scores on the criterion measures, SKOLD and SPELT, and would also (b) be found to be language disordered using BESS.

Although the SKOLD was designed for children between the ages of 2;6 and 4;0, all subjects were administered the SKOLD as an additional criterion measure, using the numerical scores rather than the pass/fail criterion since all but one of the subjects were about the upper cut off age for the SKOLD. The decision was made to use SKOLD under these circumstances because of the lack of available instruments that have been designed for use in differential diagnosis of Black English speaking children. It was found that all but three subjects (14) passed the SKOLD; however, none of the subjects earned the ceiling score. One might have expected that the subjects would have passed the SKOLD because they were all (with the exception of one), above the age limit of the SKOLD.
However, it would have been difficult to predict that they would have passed the SKOLD due to their known language disorder. Of the three subjects that failed (i.e., they scored one standard deviation or more below the mean) the SKOLD, one was within the age limit of the SKOLD (3;8), one was eight months above the age limit of the SKOLD (4;8), and one was two years; five months above the age limit of the SKOLD (6;5). These children's primary difficulties were with vocabulary, story completion, sentence repetition and comprehension of commands. Story completion, sentence repetition and comprehension of commands all require semantically based auditory memory skills. These children demonstrated that they may have difficulty with auditory memory inasmuch as they were able to complete, repeat, and follow through only on parts of stories, sentences and commands. The subjects that failed the SKOLD were also found to be language disordered on the SPELT and BESS. Their primary difficulties on the SPELT were with the present progressive ending (ing) and possessive pronouns. Their difficulties with the BESS included problems with word order and past tense.

Sixteen of the seventeen subjects scored below suggested cut off scores on the SPELT. The child who was found to be language disordered on the BESS but not on the SKOLD scored at the cut off score on the SPELT. This child was the only subject not identified as language disordered on the SPELT. It is difficult to determine why only one child passed the SPELT, but perhaps this child had progressed in speech-language therapy prior to testing for this study. Based on this information, 14 subjects did not fail the SKOLD and one subject did not
fail the SPELT as predicted by Hypothesis One (a). Although Hypothesis One predicted that all subjects would be identified as language disordered with the SKOLD, it would have been difficult to predict that the subjects would pass the SKOLD as discussed previously. A high percentage of the subjects did score above the cut off score for the SKOLD (82 percent) and below the cut off score for SPELT (94 percent).

It was noted that BESS gave credit for more Black English features than the SKOLD or SPELT did. These features included: irregular past tense; participle; negatives; yes/no question interrogatives; deletion of do in yes/no interrogatives; and WH questions. For example, BESS gave credit for variations in irregular past tense forms (e.g., He find the money yesterday/"He found the money yesterday"), where SPELT and SKOLD did not. This information indicates that BESS provided more opportunity for the child to use Black English features without being penalized for their use than did the SKOLD or SPELT.

Three subjects were not identified as language disordered by the BESS (i.e., they scored within one standard deviation of the mean). These three subjects were found to miss items on the SPELT (which found them to be language disordered) that were used appropriately on the BESS. This finding indicates that (1) the subjects may have had more opportunity on the BESS to use their own structures (Vaughn-Cooke, 1983) and/or, (2) the subjects may have misunderstood some of the examiner's instructions on the SPELT.

These findings would seem to lead to the rejection of Hypothesis One (b) since not all subjects failed the BESS. However, only a small
percentage (18 percent; 3) of the subjects were not found to be language disordered using BESS. This fact indicates that for this sample there is an 82 percent chance of BESS identifying language disordered children as being language disordered. It also illustrates the need for children to be assessed with multiple measures before determining them to be speech and language disordered. No one test is sufficient for making such judgments.

Hypothesis Two

Hypothesis Two was that there would be a significant rank-difference correlation coefficient between scores on the BESS and the criterion measures (SKOLD and SPELT). The rank-difference correlation coefficient method revealed a high correlation coefficient of .997 between the SPELT and BESS and a correlation coefficient of .677 between the SKOLD and BESS.

One of the major problems in attempting to establish the concurrent validity of an evaluation tool is the selection of appropriate criterion tests with which to compare it (Vaughn-Cooke, 1983). The SPELT was standardized on Caucasian children yet provided alternate responses for speakers of Black English dialect. The SKOLD was standardized on Caucasian and Black children, separately, who spoke Standard English and Black English respectively. Each test (Standard and Black English) receives the same credit for appropriate responses. Vaughn-Cooke (1983, p. 30) reports that, "any revision which does not provide equal credit for comparable non-mainstream and Standard English forms is inadequate". These results would appear to indicate that...
the BESS is also a valid tool for evaluation of speakers of Black English who have language disorders.

Based on these data, Hypothesis Two has been supported; there was a significant rank-difference correlation between the BESS and the criterion measures.

**Hypothesis Three**

Hypothesis Three was that a significant positive correlation (at the .05 level of confidence) would be found between (a) BESS and SPELT, and (b) BESS and SKOLD scores, using Pearson product moment correlation methods.

Pearson product moment correlation coefficient revealed a relatively high $r$ value of .81 between the BESS and SPELT, but a less strong, though significant correlation ($r$ value of .66) between the BESS and SKOLD. This method was also used to determine correlation between SPELT and SKOLD. A high $r$ value of .929 was found. This result would seem to indicate that the SPELT and SKOLD are very similar tests, however, they are not. Perhaps the strong correlation between the SKOLD and SPELT is due to the fact that they are both structured (formal) tests (i.e., specific questions and responses required by the examiner and the subject), where the BESS is less formal. The SPELT and the BESS are more similar in the language structures that are tested (see Table 1), than either of them is to the SKOLD. It would seem that the most similar sections of the SKOLD, as it relates to the BESS, was the "paired" and "individual sentence repetition" tasks. Sentence repetition tasks are based on the premise
that children repeating sentences interpret what they hear and reproduce it according to the language skills they possess. In this section the SKOLD tests the child's syntactic sequencing and comprehension abilities. It is, therefore, not surprising that BESS and SPELT were more highly correlated than the BESS and SKOLD.

Because the Pearson product moment correlation method is a more powerful correlation method, it should be viewed as a better indicator of the validity of the BESS. However, one must also keep in mind that the number of subjects was relatively low for using Pearson product moment correlation and that one of the basic assumptions of the SKOLD was violated when it was administered to children above the ceiling age of the test. It is not known exactly what the effect is of administering such a screening tool to children outside of the ages for which it was validated would be, but it is less likely to be a sensitive indicator of disorder when used in this manner. One should keep these limitations in mind and be cautious in the interpretation of results. However, the results from the Pearson product moment correlation method support both parts (a) and (b) of Hypothesis Three at the .05 level of confidence.

**Hypothesis Four**

Hypothesis Four was that the percent of interscorer agreement between two examiners on four randomly selected BESS transcripts would be high (90 percent or greater).

There was 90 percent agreement in scoring the BESS between the examiners. This percentage indicates high interscorer reliability.
which implies that several speech-language pathologists who have been prepared to use BESS should be able to gather a 50 sentence spontaneous language sample and obtain appropriate results with it. However, because BESS depends more on clinician determination of items to be scored and assignment of values, this should be done only after a training phase and practice in scoring the BESS.

The training phase consisted of selecting three transcripts from the normative study (Nelson, 1983) which were previously scored. Using the originals as a key, they were re-scored by Y.H. The investigator found that it took an hour to score one transcript using BESS/DSS. After scoring five transcripts, Y.H. found that it took only 13 and one-half minutes to complete each BESS and DSS separately for a total of 27 minutes.

Implications

Several implications regarding the BESS and other issues related to dialectally different children have emerged from this study. The first implication is that there may be hesitancy to label children at young ages (3;0 - 4;0) as language disordered within some school systems. While there is no firm agreement in the profession on how best to serve young children, systems may need to take care that efforts to avoid cultural bias in testing and overcompensation does not occur. This implication was not drawn from the results of this study, but was based on the difficulty the investigator faced in attempting to identify subjects for the study.

The second implication of this study is that there is a lack of
appropriate language test instruments for children who are culturally different; therefore, it was difficult to identify appropriate concurrent measures for this study. SKOLD was normed on standard and Black English speaking children; however, separate scores were reported and averaged together to obtain the standardization data. The SPELT was standardized on Standard English speaking children and alternative responses for the Black English speaking population were provided. The SKOLD and SPELT were validated with instruments that are not validated on a population that speaks Black English dialect. Nevertheless, they were chosen as the most appropriate measures available for use as a concurrent measure for this study. Although these tests (SKOLD and SPELT) were not specifically validated on the Black English speaking population, the high correlation between these measures and the BESS may further support their (SKOLD and SPELT) validity on dialectally different children. This is a circular argument and should be interpreted with caution.

The third implication is that a training and practice phase is necessary before one is able to accurately and quickly score the BESS. The investigator found that the average time to score each test was 27 minutes after the training and practice phase.

The fourth implication of this study is that the high percentage of interscorer agreement of scores indicate that speech-language pathologists, who have a background in language development, should be able to accurately administer and score the BESS after training and practice with the BESS.

The fifth implication is that all tests used during this study
focused on different aspects of language. The SKOLD focused on vocabulary, sentence repetition and comprehension of commands; the SPELT focused on morphology and syntax; and the BESS focused on language form (primarily morphology). The BESS and SPELT test similar structures but use different methods of elicitation. The SKOLD would seem to be more likely to miss children who might be language disordered due to the structure of the test. Although sentence repetition and comprehension of commands sections of the SKOLD require semantically based auditory memory, some children are quite able to repeat complete and complex sentences without understanding them. This process could aid in passing the SKOLD even if the child was, in fact, language disordered. The SPELT would seem to be more likely to over-identify children who might have language disorders due to the scoring format of the test. There are several Black English features that are not credited on the SPELT, as mentioned previously (see page 42 of text).

The final implication of this study is that although BESS appeared to be a sensitive measure where most of the subjects in this study fell at least one standard deviation below the mean, three subjects passed the BESS. There may have been several reasons for these results: (1) the subjects' language disorder may not have been attributed to the form of language which is primarily tested with the BESS; (2) these children may have improved in language therapy prior to testing for this study; and (3) the BESS may have over credited the Black English features used by giving credit for more Black English features than other tests. It is important to remember to use comprehensive assessment techniques when analyzing language skills.
of children and one must consider all the variables that may affect the results of the test (e.g., educational background of child, language development and area of language deficiency).

Limitations of the BESS

The BESS takes approximately 27 minutes to score and considerable time to elicit and transcribe the language sample. The speech-language pathologist must be knowledgeable about the culture for which this test is designed and should have knowledge of Black English features. Before use of BESS, one must have a training and practice phase for accurate results. It would probably not be used as a screening tool, but after there is already reason to suspect that a language disorder might be present. It would also not be the appropriate choice as a diagnostic tool for children whose language is so limited as to make it impossible to elicit 50 complete sentences in an hour or less. However, in those cases, the symptom of language paucity might be diagnostic in itself. Finally, BESS is only able to focus indirectly on semantic and pragmatic factors by failing to award the sentence point when such errors occur (i.e., If the sentence is correct in form, yet incorrect semantically, it does not receive a sentence point).

Since the SPELT and the SKOLD correlate with the BESS, one might ask why the BESS should be used if it requires more time and training than the other measures? Although there are a few limitations of the BESS, it provides more opportunity than the SPELT or SKOLD for a child to use Black English features and be credited for them. Therefore, fewer children will be inappropriately identified as being language
disordered, when in fact, they are not (i.e., fewer false positives). Based on this information, fewer children will be inappropriately placed in special education service areas than on other tests, and BESS is a more sensitive measure than the SKOLD for children above the age of 3;8. Since the BESS is based on spontaneous language samples, it provides, "richer" information for analyzing language. It can provide information (taken from the language sample) that may be used in designing individual language programs. This information is gathered from observing other aspects of the child's language during testing (i.e., non-verbal language; word order; how language is used; and the different subjects that the child talks about). Furthermore, through the use of BESS, speech-language pathologists can become more familiar with dialectal differences and use this information in training programs for other speech-language pathologists and educators. The use of BESS would also help the professionals become familiar with Black English dialect beyond the surface structure.

The decision to use one standard deviation below the mean for the determination of language disorder appeared justified by the results of the study and should be recommended as a technique for diagnostic use. Further study should be done on the BESS in effort to determine a definite level of identification of children with language disorders (i.e., one or two standard deviations below the mean).

Summary

This study has revealed that the majority of subjects identified as language disordered by IEPC's or comprehensive assessment were
found to be language disordered on the BESS and the SPELT. The SPELT and the BESS are highly correlated measures of language skills and the SKOLD and BESS demonstrated a less strong, but significant correlation for measurement of language skills. The results of this study also indicate that after a training and practice phase, speech-language pathologists with background in language development, should be able to accurately elicit spontaneous language samples and score BESS. Several implications have evolved from this investigation of concurrent validity and interscorer reliability of the Black-English Sentence Scoring system, which further support BESS as a useful tool for all speech-language pathologists and educators that work with culturally different children.

With these findings, it should be concluded that BESS is a valid measure for identifying culturally different children as being language disordered and has been shown to be reliable between scorers.
APPENDIX A

Definition of Terms
<table>
<thead>
<tr>
<th>Definition of Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bilingualism</strong></td>
</tr>
<tr>
<td>Bilingualism refers to the use of two or more languages by the same individual (Harrison and Trabusso, 1976; McCollum, 1981). This differs from Black English, in that, bilingualism refers to two distinct different languages and Black English refers to a variety of one language.</td>
</tr>
<tr>
<td><strong>Black English</strong></td>
</tr>
<tr>
<td>Black English is a variety of English that differs in phonology, grammar, and vocabulary from Standard English (Johnson, 1978; Williams, 1976). Not all Black Americans speak Black English and not all individuals who speak Black English are Black (Bently and Crawford, 1973; Dale, 1972; Dillard, 1972).</td>
</tr>
<tr>
<td><strong>Children of Diverse Cultures</strong></td>
</tr>
<tr>
<td>&quot;Children of diverse cultures&quot; (p. 1 of text) in this context does not refer to bilingual children, but children who speak a variety of English known as Black English.</td>
</tr>
<tr>
<td><strong>Ebonics</strong></td>
</tr>
<tr>
<td>This term is preferred by some professionals to replace the term Black English. &quot;Derived from the words ebony</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
and phonics, the term Ebonics is intended to avoid the focus on race and emphasize the ethnolinguistic origin and evolution of this variety of the English language" (Committee on the Status of Racial Minorities, 1983, p. 23). An Alternate viewpoint is that Ebonics is considered by some professionals to be a separate system from Black English (Wofford, 1979).

**Language Difference**

Language difference refers to a "normal variation in language from community to community which is related to factors such as religion, class, ethnicity" (Wolfram, 1979, p. 1) and includes variations associated with roles, socioeconomic status, and education (Adler and Birdsong, 1983; Taylor and Payne, 1983).

**Language Disorder**

Language disorder occurs when certain forms "do not match the norms of a particular community by deviation from the community's language pattern" (Wolfram, 1979, p. 1).

**Nonbiased Assessment**

Nonbiased assessment refers to evaluation procedures administered so that
they are not racially or culturally discriminating (Anderson, Freeman and Edwards, 1977; Ballard, 1975) and evaluation procedures should be provided in the child's native language or mode of communication (Anderson, Freeman, Edwards; 1977; Ballard, 1975).

**Standard English**

Standard English is the form of English that is used for language in grammar books, dictionaries and "language academies (schools)" and is utilized by the majority of people (Wolfram and Fasold, 1974).

**Validity**

Validity is the extent to which a test measures what it is supposed to measure (Leeman, 1981).
APPENDIX B

Scoring System for BESS
<table>
<thead>
<tr>
<th>Points</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
| 3      | -and plus  
- and (when intonation makes sentence combination clear): He pointed his finger at him, (with rising intonation) he pointed his finger at him (with falling intonation). |
| 4      | -don't (with 3rd pers. sing. as 2nd or 3rd neg. marker): He don't want none. No one don't live with me.  
-can't, don't (as proposed aux): Can't nobody make me.  
-3 copula/aux = not + V: My mama not gonna pick me up today. He not a baby.  
-sin't (as negative copula or aux, he): He ain't my friend.  
-won't (as proposed aux): Don't nobody help him.  
-for/sot: The dog make too much noise for they won't catch many fish.  
-conditional and: You do that and I'm gonna smack you.  
-if with phrase deletions: He lookin' if he see the money.  
-aux inversion in indirect Qs (instead of if): She ask me do he want some more.  
-or either, or neither (as disjunctives): He will go or either he will stay. He told her that he wouldn't be bad or neither get in trouble.  
-proposed why phrase (with because): 'Why he's in here, cause I'm scared the dog.  
| 5      |  |
| 6      | -ain't (for have + not) ≠ uninflected V: I ain't taste any.  
-ain't (for did + not) ≠ uninflected V: Yesterday, he ain't go to school. I ain't found thing in the school.  
-couldn't, couldn't, shouldn't (as proposed aux)  
-shan't/shan't: The walks wasn't walkin' right.  
-won't/shan't: There wasn't no more.  
-contrary, contrary neg.  
-will/shan't: I didn't get to 1/13/14 to come in the building.  
-ain't: I've been in that store in him, and the sale.  
-won't: I didn't get to.
<table>
<thead>
<tr>
<th>Points</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

- who, what, what, when (with deleted aux or copula)
- where, how, why, how much, what, do, what...
- for (with deleted aux or copula): where the man?
- How's formed without interrogative reversal: what that is?

- aux: by voice
- gonna come out of here?
- you gonna tell my mana?
- was/were: keep throwin' rocks?

- how (with deleted aux, copula or do): be this how you do this?

- do: you know that one with the tractor? where you work? you got blue eyes?
- do (with 3rd pers. sing.) be he still have it?
- 3 or unversed modal: how, what else I be doin? why you can't talk on that?
- tag question with ain't: it gonna be fun, ain't it?

- why, what if, how come (with deleted or unversed aux, copula or do, or with get): why she turn that way? Hey, why she got a dress on here?

- reduced form: who ain't: how you want that you been doing?
- ain't with 3rd pers. sing.: who he left you?
- you, we, they (with deleted or unversed copula or to): why the hell the sky is blue?
APPENDIX C

Language Analysis Form for BESS and DSS
Subject 21
Age 4:0
Date May 31, 1984 DSS 3.18 BESS 5.90

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1,1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1,2</td>
<td>2</td>
<td>1</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1,3</td>
<td>1</td>
<td>4</td>
<td>1,2</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>3,1</td>
<td>2</td>
<td>5</td>
<td>1,2</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Total DSS for this partial sample:
35 divided by 11 = 3.18

Total BESS for this partial sample:
65 divided by 11 = 5.90
APPENDIX D

Permission Form for Subjects and Their Parents
Permission Form

Child's name ___________________________ Birthdate ___________________________

Parent name ___________________________
Mother ___________________________
Father ___________________________

Occupation ___________________________
Mother ___________________________
Father ___________________________

Education ___________________________
Mother ___________________________
Father ___________________________

Address ___________________________ Telephone ___________________________

Does your child have any difficulty talking? Yes No
If yes describe ____________________________________________________________

Has he/she had any difficulty hearing? Yes No
If yes describe ____________________________________________________________

Does your child have behavioral difficulties? Yes No
If yes describe ____________________________________________________________

The taped conversation between your child and the interviewer will be erased after the research is completed.

This study "Validity of the Black English Sentence Scoring System" has been explained to me, and I grant my permission for my child ___________________________ to participate in this project. I understand that I may withdraw my permission at any time, and that my child also has the right to decide not to participate.

Parent Signature ___________________________ Date ___________________________

This study "A Validity Study of the Black English Sentence Scoring System" has been explained to me and I do not give my permission for my child ___________________________ to participate in this project.

Parent Signature ___________________________ Date ___________________________

I would like further information about ___________________________ before making my decision.
APPENDIX E

Social Status of Parents
The parents of the subjects were classified according to their level of education and occupation. In each group, education and occupation, there are seven levels, where 1 is the highest level obtained and 7 is the lowest (Warner, 1960).

Following is a list of the education and occupational levels:

**Education**

1. Professional/Graduate School
2. College Education (1-4 years)
3. Highschool Graduate
4. One to Three Years of Highschool
5. Grammar School Graduate (eighth grade)
6. Four to Seven Years of School
7. Zero to Three Years of School

**Occupation**

1. Inherited Wealth
2. Earned Wealth (savings, investments)
3. Profits/Fees (owners of businesses)
4. Salary (monthly/yearly basis)
5. Wages (hourly rate/daily basis)
6. Private Relief (given money by friends/relatives)
7. Public relief (State assistance)
The following is a summary of the average household educational and occupational levels of the subjects' parent(s).

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average Occupation</th>
<th>Average Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>4.5</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>5.5</td>
<td>2.5</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>—*</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>2.5</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>2.5</td>
</tr>
<tr>
<td>14</td>
<td>5.5</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>4.5</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

* —* refers to unknown occupational level and/or educational level.
APPENDIX F

Raw Data Chart
<table>
<thead>
<tr>
<th>Subject-Number</th>
<th>C.A.</th>
<th>Sex</th>
<th>Geo*</th>
<th>Placement</th>
<th>BESS</th>
<th>Pass Fail</th>
<th>DSS</th>
<th>Pass Fail</th>
<th>SKOLD</th>
<th>Pass Fail</th>
<th>SPELT</th>
<th>Pass Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3;8</td>
<td>F</td>
<td>K</td>
<td>Headstart</td>
<td>5.20</td>
<td>F</td>
<td>4.42</td>
<td>F</td>
<td>10</td>
<td>F</td>
<td>10</td>
<td>F</td>
</tr>
<tr>
<td>2</td>
<td>5;5</td>
<td>M</td>
<td>K</td>
<td>&quot;</td>
<td>7.40</td>
<td>F</td>
<td>4.92</td>
<td>F</td>
<td>30</td>
<td>P</td>
<td>29</td>
<td>F</td>
</tr>
<tr>
<td>3</td>
<td>5;7</td>
<td>M</td>
<td>K</td>
<td>&quot;</td>
<td>9.28</td>
<td>F</td>
<td>5.82</td>
<td>F</td>
<td>30</td>
<td>P</td>
<td>31</td>
<td>F</td>
</tr>
<tr>
<td>4</td>
<td>4;8</td>
<td>M</td>
<td>K</td>
<td>&quot;</td>
<td>5.86</td>
<td>F</td>
<td>3.82</td>
<td>F</td>
<td>17</td>
<td>F</td>
<td>16</td>
<td>F</td>
</tr>
<tr>
<td>5</td>
<td>6;0</td>
<td>M</td>
<td>K</td>
<td>&quot;</td>
<td>8.26</td>
<td>P</td>
<td>5.96</td>
<td>F</td>
<td>25</td>
<td>P</td>
<td>29</td>
<td>F</td>
</tr>
<tr>
<td>6</td>
<td>5;5</td>
<td>M</td>
<td>K</td>
<td>Prepri-</td>
<td>9.14</td>
<td>P</td>
<td>7.12</td>
<td>P</td>
<td>34</td>
<td>P</td>
<td>39</td>
<td>F</td>
</tr>
<tr>
<td>7</td>
<td>5;11</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>6.58</td>
<td>F</td>
<td>2.60</td>
<td>F</td>
<td>25</td>
<td>P</td>
<td>23</td>
<td>F</td>
</tr>
<tr>
<td>8</td>
<td>6;1</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>7.66</td>
<td>F</td>
<td>5.34</td>
<td>F</td>
<td>35</td>
<td>P</td>
<td>38</td>
<td>F</td>
</tr>
<tr>
<td>9</td>
<td>5;7</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>6.48</td>
<td>F</td>
<td>5.22</td>
<td>F</td>
<td>34</td>
<td>P</td>
<td>30</td>
<td>F</td>
</tr>
<tr>
<td>10</td>
<td>5;6</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>6.10</td>
<td>F</td>
<td>4.70</td>
<td>F</td>
<td>30</td>
<td>P</td>
<td>27</td>
<td>F</td>
</tr>
<tr>
<td>11</td>
<td>4;9</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>4.61</td>
<td>F</td>
<td>3.18</td>
<td>F</td>
<td>21</td>
<td>P</td>
<td>18</td>
<td>P</td>
</tr>
<tr>
<td>12</td>
<td>6;2</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>6.66</td>
<td>F</td>
<td>5.42</td>
<td>F</td>
<td>27</td>
<td>P</td>
<td>32</td>
<td>F</td>
</tr>
<tr>
<td>13</td>
<td>6;0</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>6164</td>
<td>F</td>
<td>5164</td>
<td>F</td>
<td>29</td>
<td>P</td>
<td>20</td>
<td>F</td>
</tr>
<tr>
<td>14</td>
<td>6;2</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>7.22</td>
<td>F</td>
<td>5.66</td>
<td>F</td>
<td>35</td>
<td>P</td>
<td>34</td>
<td>F</td>
</tr>
<tr>
<td>15</td>
<td>5;7</td>
<td>F</td>
<td>D</td>
<td>&quot;</td>
<td>5.70</td>
<td>F</td>
<td>4.62</td>
<td>F</td>
<td>19</td>
<td>P</td>
<td>14</td>
<td>F</td>
</tr>
<tr>
<td>16</td>
<td>5;6</td>
<td>M</td>
<td>D</td>
<td>&quot;</td>
<td>5.18</td>
<td>F</td>
<td>3.42</td>
<td>F</td>
<td>26</td>
<td>P</td>
<td>14</td>
<td>F</td>
</tr>
<tr>
<td>17</td>
<td>6;5</td>
<td>F</td>
<td>D</td>
<td>&quot;</td>
<td>5.28</td>
<td>F</td>
<td>3.30</td>
<td>F</td>
<td>16</td>
<td>F</td>
<td>14</td>
<td>F</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


Committee on CCC Language State (1975). Student's rights to their own language. College English, 36, 709-726.


