

COMPARISON BETWEEN A DYNAMIC ASSESSMENT PROCEDURE AND THE
WMLS-R IN DISTINGUISHING LANGUAGE ABILITIES AMONG HISPANIC
CHILDREN IN FIRST GRADE
A DISSERTATION
SUBMITTED TO THE FACULTY
OF
THE GRADUATE SCHOOL OF APPLIED AND PROFESSIONAL PSYCHOLOGY
OF
RUTGERS,
THE STATE UNIVERSITY OF NEW JERSEY
BY
JILL MARIE SPERO
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
OF
DOCTOR OF PSYCHOLOGY

NEW BRUNSWICK, NEW JERSEY

MAY 2012

APPROVED:

Geraldine V. Oades-Sese, Ph.D.

Karen L. Haboush, Psy.D.

DEAN:

Stanley B. Messer, Ph.D.

ABSTRACT

Historically, educators have had significant difficulty assessing the needs of culturally and linguistically diverse (CLD) learners, especially when determining special education classification. Hispanic students seem especially vulnerable to schools' traditionally inadequate means of assessing language ability in CLD students. Dynamic assessment is one approach that has been identified as promising in the development of more culturally competent evaluation procedures. The purpose of this dissertation was to compare a traditional language assessment to a dynamic assessment procedure in regard to their utility for identifying students at-risk for Speech or Language Impairment (SLI). Twenty-five Hispanic, bilingual first graders from an urban school district were administered a standardized measure of language (Woodcock-Munoz Language Survey-Revised) as a traditional indicator of language ability. Students scoring one standard deviation below the mean were labeled at risk for SLI classification, while those scoring at the mean or higher were deemed typically developing. All children then underwent a dynamic assessment of language ability involving a pretest, two 30-minute mediated learning experiences, and a posttest. Following dynamic assessment, both groups made significant improvements in story complexity and completeness, and did so to a comparable degree. Furthermore, 100% of students classified at risk for SLI by the standardized measure were deemed typically-developing according to dynamic assessment. Consistent with previous dynamic assessment studies, the results indicate that reliance on a traditional, standardized measure of language to assess Hispanic students for SLI may underestimate their language abilities and, thus, increase the risk for special education classification. Further, dynamic assessment appears to be a valuable tool for discerning language

differences from true language disorders in Hispanic children. Both dynamic assessment and standardized measures of language may assess different aspects of language ability, which together may comprise a more reliable assessment strategy for identifying SLI in Hispanic children. Implications for the use of dynamic assessment in SLI evaluation and intervention are discussed in light of concerns about misclassification of CLD students. Additionally, implications for the role of school psychologists related to dynamic assessment are addressed.

ACKNOWLEDGEMENTS

Thank you to Gerry and Karen for your generous support throughout the dissertation process. Your thoughtful guidance and feedback at every step was invaluable. Also, thank you to the Resilience Lab for all of your help during this project.

Thank you to Elena and BJ for being my dissertation support group and for your willingness to help me sort through things whenever I got stuck. And to Elena and Victoria for your emotional support and friendship. Thank you for the trips to Taco Bell, the bad TV, and keeping Nutella in your houses for me.

Thank you to my cohort for being different. There is no other group of people who I would've wanted to go through these last three years with.

Thank you to the Albro family for giving me a second home and for the bagels every Saturday morning.

Thank you to Michael and my family for everything.

TABLE OF CONTENTS

	PAGE
ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES.....	ix
CHAPTER	
I. INTRODUCTION.....	1
Background and Context.....	1
Research Questions and Hypotheses.....	4
Significance of the Study	6
Definitions of Terms	6
II. LITERATURE REVIEW.....	8
Disproportionate Representation in Special Education of Culturally Diverse Students.....	8
Hispanic Students in Special Education.....	10
Theories of Language Development	11
Components of Language	14
Stages of Language Development.....	15
Second Language Acquisition and Bilingualism	17
Language Demands in School.....	20
Speech or Language Impairment (SLI).....	22

Assessment Procedures for SLI	23
Dynamic Assessment	28
III. METHOD.....	36
Participants	36
Measures	36
Procedure.....	37
Scoring Procedures.....	39
Data Analysis Procedures	40
IV. RESULTS	42
Zero Order Correlations	42
Oral Language and English Proficiency.....	42
Research Question 1	44
Research Question 2.....	46
Research Question 3	49
V. DISCUSSION	54
Language Abilities According to a Standardized Measure of Language	55
Language Abilities According to Dynamic Assessment.....	56
Traditional Assessment Procedures Versus Dynamic Assessment.....	58
Implications for the Use of Dynamic Assessment in SLI Classification Procedures	60
Implications for School Psychologists	63

Implications for the Use of Dynamic Assessment in Developing	
Interventions.....	65
Limitations	70
Future Research and Conclusions	73
REFERENCES.....	75
APPENDICES.....	84

LIST OF TABLES

Table 1. Zero Order Correlations Among Variables of Interest	43
Table 2. Descriptive Statistics for Typically-Developing Children and Children At Risk for SLI Classification	44
Table 3. Descriptive Statistics for Total Story and Productivity at Pretest and Posttest for Typically-Developing Children and Children At risk for SLI Classification	47
Table 4. Descriptive Statistics of Gain Scores for Typically-Developing Children and Children At Risk for SLI Classification.....	49
Table 5. Descriptive Statistics of Modifiability Ratings for Typically-Developing Children and Children At Risk for SLI Classification	49

LIST OF FIGURES

Figure 1. Growth in Number of Words for Children At Risk for SLI Classification 48

Figure 2. Growth in Number of Words for Typically-Developing Children 48

CHAPTER I

INTRODUCTION

Background and Context

One of the most important foundations of education is assessment of students' strengths and needs. Satisfactory assessment allows for a comprehensive understanding of children, which thereby ensures that they are provided the educational programming they need. Thus, they are afforded the opportunity to benefit from public education. Conversely, unsound assessment procedures may greatly inhibit children's access to necessary educational supports, limit their educational growth, and contribute to long-term negative outcomes. Therefore, it is important for education professionals to develop competent assessment procedures in order to best promote school success and lasting success for children. One arena where this concern has been focused is the determination of special education eligibility. In this case, assessment not only establishes whether or not a child qualifies for special education services, but also directs his or her educational placement and instructional programming. The implications are significant: It is critical that educators accurately evaluate and classify students for special education because these determinations have pervasive, enduring effects on children.

Historically, educators have had a particularly difficult time accurately assessing culturally and linguistically diverse (CLD) students for special education within a variety of classification categories. This has resulted in the disproportionate placement of CLD

students in special education, which has been seen as a discriminatory practice against a historically marginalized group (Marbley, Bonner, & Berg, 2008). Furthermore, inappropriate special education assessment has been associated with negative consequences for CLD students, including denied access to necessary education services, restricted contact with general education programming, and academic underperformance (Patton, 1998; Perez, Skiba, & Chung, 2008; Poon-McBrayer & Garcia, 2000; Reid & Knight, 2006). The literature suggests that developing sound, culturally competent assessment procedures and training culturally competent professionals is critical to achieving sound special education assessments for CLD students (Marbley, Bonner, & Berg, 2008; Schon, Shaftel, & Markham, 2008; Skiba, Knesting, & Bush, 2002).

Hispanic children are one such important CLD group that should be considered when discussing culturally competent special education assessment, both due to their growing number within the United States population and their unique needs. The number of Hispanic children enrolled in public schools in the United States is growing rapidly and is projected to continue growing in the next 40 years (Pew Hispanic Center, 2008). There are now ten million Hispanic students in American public schools, or approximately one-in-five students. By 2050, the U.S. Census projects that this number will increase by 166 percent to roughly 28 million. In contrast, the enrollment of non-Hispanic children in public schools is projected to increase by only 4 percent. Further, Hispanic students are more likely than their non-Hispanic peers to live in poverty and have parents who did not finish high school. Also significant is the language background of the Hispanic student population. Specifically, about 70 percent of Hispanic students speak a language other than English at home (most commonly Spanish), while 30 percent

speaking only English at home. Thus, the majority of Hispanic children in American schools are bilingual. Additionally, while first-generation Hispanic students have the most difficulty with the English language, second, third, and higher generation children can also lack proficiency in English: 44 percent of first-generation, 20 percent of second-generation, and 5 percent of third-and-higher generation Hispanic students speak English with difficulty (Pew Hispanic Center, 2008). These facts point toward a demand for educators to be increasingly sensitive to Hispanic students, who will likely enter school with a unique set of needs. Of note, the term “Hispanic” will be used for the remainder of this document in order to correspond with the terminology most frequently utilized in the literature. However, it is recognized that this terminology is persistently evolving and that many are in favor of using other terms, including “Latino” or more specific indicators of national origin (e.g., “Mexican,” “Colombian”).

In light of these factors and considering the historic disproportionate placement of CLD students in special education, it is relevant to examine the current procedures for disability evaluation and classification of Hispanic children. It seems imperative to consider this process in relation to the classification of Speech or Language Impairment (SLI), as Hispanic students may be distinctly susceptible to the procedural inefficiencies in the evaluation of this classification (Harry, 1994). In particular, the standardized measures typically used in schools to assess SLI appear to be ill equipped to delineate language differences from language disorders in Hispanic students (Gillam, Pena, & Miller, 1999; Marbley, Bonner, & Berg, 2008; Miller, Gillam, & Pena, 2001; Shames & Anderson, 2002). This may then result in Hispanic children being classified with SLI

when a language disorder does not exist or, conversely, Hispanic children not being found eligible for SLI when they truly have a language disorder.

Dynamic assessment is a procedure that is thought to be a promising means of distinguishing language difference from a language disorder, especially in CLD students (Guitierrez-Clellen & Pena, 2001; Haywood & Lidz, 2007; Lidz, 1991; Pena, Iglesias, & Lidz, 2001; Pena et al., 2006). Dynamic assessment is based upon Vygotsky's (1978, 1986) zone of proximal development (ZPD) and provides a representation of ability based upon the individual's ability to improve in a given area with guidance from a competent teacher (Haywood & Lidz, 2007). Because dynamic assessment relies not on standardized measures, but on one's ability to learn, it is thought to be less susceptible to environmental and experiential biases than traditional assessment procedures. As a result, it has been recognized as a potentially more culturally competent means of evaluating CLD children for SLI. In particular, the gains the child makes during dynamic assessment and the nature of the intervention required to produce these gains (i.e., the child's modifiability) appear to be important in making SLI classification decisions (Haywood & Lidz, 2007; Lidz, 1991; Miller, Gillam, & Pena, 2001; Pena, Iglesias, & Lidz, 2001; Pena, Queinn, & Iglesias, 1992).

Research Questions and Hypotheses

The purpose of the current study is to compare a traditional standardized measure of language and a dynamic assessment procedure for identifying children at risk for SLI classification in a group of 25 Hispanic, bilingual first grade students.

The following research questions were examined in this study:

1. Do Hispanic, bilingual first graders make significant improvements in measures of their language abilities during a dynamic assessment procedure?
2. For Hispanic, bilingual first graders, who may be at risk for special education classification under the category Speech or Language Impairment based on standardized measures of language: Do these students perform differently than typically-developing children in a dynamic assessment procedure of their language abilities?
 - a. Do they perform differently on ratings of modifiability?
3. Based on this pilot study, is dynamic assessment of language ability a valuable addition to the Speech or Language Impairment assessment process, especially for culturally and linguistically diverse students?

The hypotheses are as follows:

1. Hispanic, bilingual first graders will make significant improvements in measures of their language abilities during a dynamic assessment procedure, since the method allows for a “test-teach-retest” procedure.
2. Hispanic, bilingual first graders, with typically-developing language abilities based on standardized measures of language, will make significantly more gains than children at risk for Speech or Language Impairment classification in a dynamic assessment procedure of their language abilities.
 - a. They will also perform significantly better than children at risk for Speech or Language Impairment classification on ratings of modifiability.

3. Dynamic assessment provides valuable information about language learning ability and, thus, may be a culturally sensitive addition to the Speech or Language Impairment assessment process.

Significance of the Study

The current study will examine two approaches to assessing language ability in Hispanic, bilingual first grade students: a traditional standardized measure and a dynamic assessment procedure. Due to the apparent shortcomings of assessing CLD children with traditional assessment procedures, it is important to investigate novel techniques for doing so that may be more culturally sensitive. Accurately assessing CLD students is especially critical in order to provide them with the appropriate educational services. The current study aims to examine the utility of incorporating a dynamic assessment procedure into the SLI assessment process. Previous studies have used dynamic assessment to examine group differences in language ability among students with and without language impairment, many of whom were from culturally and linguistically diverse backgrounds. However, research of within-group differences among Hispanic bilinguals is limited.

Definitions of Terms

Bilingual: individuals who possess various levels of oral language proficiency in English and a second language (i.e., Spanish).

Disproportionate Representation in Special Education: the extent to which membership in a given ethnic group affects the probability of being placed in a specific special education disability category.

Dynamic Assessment: An assessment approach in which the evaluator provides mediation of cognitive processes in order to gauge the individual's responsiveness to intervention and learning potential.

Expressive Language: the ability to produce language.

Language Impairment: the inability to learn language as manifested by deficits in expressive and/or receptive language skills relative to age-matched peers who have comparable language exposure.

Mediated Learning Experience: in a dynamic assessment procedure, the process by which the evaluator interacts with an individual in order to produce improvement in a particular domain.

Modifiability: the nature and intensity of the intervention necessary to affect change in the individual during a dynamic assessment procedure.

Receptive Language: the ability to comprehend language.

Speech or Language Impairment (SLI): as defined in the Individuals with Disabilities Education Improvement Act (2004), a communication disorder, such as stuttering, impaired articulation, a language impairment, or a voice impairment that adversely affects a child's educational performance.

CHAPTER II

LITERATURE REVIEW

Disproportionate Representation in Special Education of Culturally Diverse Students

The disproportionate representation of culturally diverse students in special education has been an enduring focus in the field for many years due to the associated negative outcomes for these students (Fletcher & Navarrete, 2003; Skiba et al., 2008). MacMillan and Reschly (1998) propose that the proportion of particular ethnic groups eligible for special education within any classification category should equal the proportion of that group within the school population. When this is not the case, it has been suggested that contextual factors (e.g., inappropriate classification procedures) contribute in some way; disability evaluation practices that use cultural norms to define what is and is not “normal” appear to be particularly sensitive to these factors (MacMillan & Reschly, 1998; Reid & Knight, 2006). Oswald, Coutinho, Best, and Singh (1999) define the degree of disproportionate representation as “the extent to which membership in a given ethnic group affects the probability of being placed in a specific special education disability category” (p. 198). As such, a group is said to be overrepresented in special education if being a member of that group increases the probability of being found eligible for special education services. Conversely, a group is said to be underrepresented in special education if belonging to that group decreases the chances that one will be found eligible for special education services.

Overrepresentation in special education is one way in which disproportionality is manifested. While special education classification is often a positive step in meeting children's needs, it can be problematic. This is especially true if a child is classified when he or she does not actually have a disability and/or if classification leads to negative outcomes for the child (Hosp & Reschly, 2003). Students placed in special education who do not truly have a disability are at risk of not receiving the services necessary for them to benefit from their educational experiences. In these cases, special education services are provided on the basis of an incorrect assessment of needs (Gottlieb, Alter, Gottlieb, & Wishner, 1994). Moreover, the U.S. Department of Education (2006) reported that about half of the nation's school-aged children with disabilities are educated in general education classrooms for less than eighty percent of the school day. This suggests that being classified with a disability results in more restrictive educational placements for about half of the children found eligible for special education services. Consequently, these students are less likely to participate in general education classrooms with their typically-developing peers and, thus, are less likely to receive the academic and social benefits of general education (Patton, 1998). Also, literature reviews indicate that the efficacy of special education programs are mixed (Reid & Knight, 2006). Therefore, students who are classified as eligible for special education and placed in such programs may be receiving subpar educational services, possibly resulting in academic underperformance, underrepresentation in post-secondary schooling, and other negative outcomes. (Perez, Skiba, & Chung, 2008; Reid & Knight, 2006).

Though, historically, the issue of overrepresentation has been fervently examined (e.g., African American students classified with intellectual disability),

underrepresentation in special education classifications is also a concern. It has been proposed that underrepresentation can result in denied access to necessary educational services (Poon-McBrayer & Garcia, 2000). This may be especially likely to occur if students are receiving increased educational support through another avenue (Guidberson, 2009). For example, Hispanic students may be evaluated for special education only after depleting English Language Learner (ELL) services, which may be a problem for those students whose needs are truly special education related (Artiles, Rueda, Salazar, & Higaeda, 2002). This results in a prolonged period during which some students' educational needs are not being met, which may lead to negative outcomes. In particular, Perez, Skiba, and Chung (2008) hypothesize that the national academic underperformance of Hispanic students may be related to schools' inability to meet the needs of their diverse learners. Therefore, one way that schools are not meeting the needs of such students may be their difficulty in identifying Hispanic students who have disabilities.

Hispanic Students in Special Education

Due to the possible negative ramifications, the Individuals with Disabilities Education Improvement Act (2004) mandates that states monitor the representation of cultural groups in special education and enact policies to prevent disproportionality. The literature examining the disproportionate representation of Hispanic students in special education presents an unclear picture of current trends in this regard. Data appear to vary over time rather than remain somewhat stable, as is true for other cultural groups (Artiles & Trent, 1994; Skiba et al., 2008). In Guidberson's (2009) review of the literature on Hispanic students in special education, he determined that their representation in special

education appeared to vary with disability classification. Hispanic students were more likely to be classified with a learning disability or speech-language impairment and less likely to be classified with an intellectual disability. Further, the most current national data indicate that, overall, Hispanic students are underrepresented in special education, including the category of Speech-Language Impairment (National Center on Culturally Responsive Educational Systems, NCCRESt, 2006; U.S. Department of Education, 2006). However, state and district-level data have also shown patterns of overrepresentation, both in special education and, more specifically, in the classification of Speech or Language Impairment (Artiles, Rueda, Salazar, & Higareda, 2002; NCCRESt, 2006; New York University, School of Education, 1994). Taken together, the variability across time, location, and disability classification suggest that disproportionality may be, in part, a function of changes in national demographics (i.e., growth of the Hispanic population in the U.S.), the increasing number of educational service options for Hispanic students, such as ELL programs, and the nature of assessment procedures (Artiles & Trent, 1994; Kummerer, 2010; MacMillan & Reschly, 1998). The difficulty establishing a concrete understanding of Hispanic representation in special education calls attention to the necessity of assessment procedures that are better tailored to the needs of Hispanic students, especially in the category of Speech or Language Impairment (Guidberson, 2009; Kummerer, 2010).

Theories of Language Development

Language development is an area of such complexity that there does not exist a single model that is readily agreed upon. Rather, four theories have emerged as principal in the field: behavioral, psycholinguistic-syntactic, psycholinguistic-semantic/cognitive,

and sociolinguistic. Owens (2005) suggests that, while these theories may all contend that they produce the most logical portrayal of language development, different aspects of each are typically drawn upon to explain relevant pieces; cases can be made both for and against each theory.

The behavioral model of language development was predominantly supported by B.F. Skinner (1957). This theory asserts that language is a learned behavior that is subject to the principles of operant conditioning: reinforcement, punishment, and extinction. Specifically, children acquire language as a result of their interaction with the environment, including the consequences provided to their verbal behavior and the verbal behavior modeled by those around them. Skinner asserted that language develops “when relatively unpatterned vocalizations, selectively reinforced, gradually assume forms which produce appropriate consequences” (p. 31). Thus, language is a series of associations between word and meaning, word and phoneme, and statement and response (Owens, 2005) that is developed through one’s interaction with the environment.

The syntactic model of psycholinguistic theory was most readily supported by Noam Chomsky (1975, 1980). This model suggests that language development is biologically based and, as such, human beings are innately programmed to acquire language. He argued that this was evident due to the consistency in core principles of language across different languages and cultural groups. According to this theory, children acquire language by developing hypotheses about basic linguistic rules and then testing them within the natural environment. These basic rules are thought to be contained within the language acquisition device (LAD), which is a mechanism thought

to innately exist in human beings. It is the LAD that allows children to manage language and to develop testable language hypotheses.

The semantic/cognitive model of psycholinguistic theory began gaining attention in the 1970s, especially with the work of Lois Bloom (1970). It was then that scholars began to argue that theories of language development must consider children's perceived semantic relationships when discussing language acquisition. According to this theory, children's first linguistic productions are demonstrations of these semantic relationships, which represent their understanding of the relationships present in the world around them. Further, these perceptions about meaning represent children's general progression of cognitive development. Therefore, as children grow cognitively, they are able to further develop their understanding of semantic relationships and, thus, their language progresses accordingly. In the semantic/cognitive model, language is the channel through which children express their knowledge of the world (McLean & Snyder-McLean, 1978; Owens, 2005).

The sociolinguistic theory of language development focuses on the social/communicative circumstances and intent surrounding language. According to this theory, it is this desire to communicate successfully that drives humans to acquire language. In order to communicate effectively, one decides on the linguistic form and content based upon one's perception of the communicative partner and the context of the communicative exchange. Then, in language development, it is the interaction with those around them that allows young children to decide what utterances will best allow for the expression of their intentions. Thus, language development is transactional in nature in that it involves child-caregiver give-and-take that allows the children to acquire basic

rules of dialogue, rather than of syntax or semantics. It is the social context that lays the necessary structure upon which children learn to understand and express language (McLean & Snyder-McLean, 1978; Owens, 2005).

Components of Language

Form, content, and use are the three principal components that make up language (Bloom & Lahey, 1978). Form is composed of syntax, morphology, and phonology, which are the elements that connect sounds and symbols to produce meaning. Content refers to semantics, or the meaning of language. The use of language encompasses pragmatics. Syntax, morphology, phonology, semantics, and pragmatics form the basic rule framework for language.

Syntax concerns the assemblage of words to produce meaningful sentences; it refers to the form or structure of sentences. As such, rules of syntax govern sentence organization and word, phrase, and clause order. Morphology refers to how words are arranged internally. A morpheme is the smallest grammatical unit that is linguistically meaningful and can be free or bound. Free morphemes can operate independently (e.g., dog, ball, sad), while bound morphemes cannot stand alone (e.g., -s, -ed, pre-). They must be joined with free morphemes and represent grammatical markers that indicate changes in meaning. Phonology is concerned with the organization of phonemes, which are the smallest units of sound in language (e.g., the /r/ sound in *red*). Rules of phonology dictate the order and distribution of phonemes. Semantics deals with how language form is connected to meaning in language. Therefore, these rules shape how words and word combinations represent ideas about objects, events, and relationships. Finally, pragmatic rules govern how to utilize language in a communicative context. They include standards

regarding organization and coherence of conversation, error repair, speaker roles, and speech acts (i.e., speaker intentions within the communicative circumstance). Pragmatics concern how one uses language in order to convey meaningful messages to others in social environments.

Stages of Language Development

Language development begins very early in life. Receptive language develops far prior to expressive language, as babies are able to understand language before they are actually able to produce words (Fenson et al., 1994). In early infancy, babies recognize and process speech sounds differently than they do other sounds (Trawick-Smith, 2003). Specifically, they can differentiate between some consonant and vowel sounds (Jusczyk, 1995; Sansavini, Bertoni, & Giovanelli, 1997) and they can attend to differences in intonation and pitch (Katz, Cohn, & Moore, 1996). Babies first start to comprehend words sometime after six months of age (Fenson et al., 1994), usually object labels (Trawick-Smith, 2003). The first words a young child understands are also dependent upon the culture within which the child is raised; different cultures place varying emphases on specific kinds of words (Trawick-Smith, 2003).

Expressive language, or productive communication, develops as babies begin communicating messages to those around them (Trawick-Smith, 2003). Early language production begins with crying, then other noises and gestures. Babbling is recurring vocalization of speech sounds and is present during a baby's first year. At four to five months of age, babbling becomes more complex and is increasingly shaped by the speech of others. Though babies do not seem to babble to articulate specific needs or ideas, babbling is important for language development because it prompts the beginnings of

conversation between babies and their caregivers. First words are typically spoken between 8 and 18 months of age. During this stage of development, children are apt to make either overgeneralization or overrestriction errors when utilizing language. Babies overgeneralize when they use one word to represent more than it should (e.g., using “dad” to mean all adult males). Overrestriction errors occur when babies use a word to correspond to fewer things or ideas than it should (e.g., understanding “cup” as referring to only his blue cup). These errors gradually decrease as children gain exposure to language. Further, babies at this stage learn two different categories of words: referential and expressive (Gleitman & Gleitman, 1992). Referential words refer to object labels, while expressive words are social communications. Which category of words children acquire more quickly is dependent upon temperament and the environment and/or culture within which children live. Between 18 and 24 months of age, children start speaking in short sentences, beginning with two-word utterances (Trawick-Smith, 2003). It is during this stage that toddlers must first make decisions regarding word order that affect sentence meaning. From this point on, children quickly learn to string together more words in to meaningful sentences in order to express more complex ideas.

During the preschool years, children’s speech becomes more fluent and easier to understand (Trawick-Smith, 2003). Children of this age also rapidly acquire new vocabulary and are able to comprehend word meaning more completely. Moreover, they become increasingly adept at utilizing language in social contexts. By age six, children are competent language users. They have acquired most of their language’s speech sounds and continue to build their vocabulary. Though some syntactic rules remain

difficult to comprehend and use at this time, they are generally able to formulate and interpret complex sentences and ideas.

Second Language Acquisition and Bilingualism

By and large, second language acquisition by young children parallels first language development (Owens, 2005). Similar theories of language development exist and children appear to follow similar developmental patterns in second language acquisition. Specifically, five stages of second language acquisition have been named: Preproduction, Early Production, Speech Emergence, Intermediate Language Proficiency, and Advanced Language Proficiency (Hearne, 2000; Northwest Regional Educational Laboratory, 2003; Rhodes, Ochoa, & Ortiz, 2005; Roseberry-McKibbin, 2002). The preproduction stage commonly includes a silent period during which the child's priority is on language comprehension; he or she might know about 500 words receptively. At this time, the child might respond to language initiatives using gestures, engaging in a relevant action (e.g., sitting down when told, "sit down"), or providing one-word answers. Children typically remain in this stage of language acquisition for their first three months of exposure to the second language. The early production stage takes place from about month three to month six of second language exposure. During this stage, children acquiring a second language continue to focus on language comprehension. They can understand and utilize about 1000 words comfortably. Children in this stage typically use one- to three-word phrases and can provide short responses to yes/no, either/or, and who/what/where questions.

From six months to two years of second language exposure, children are typically in the speech emergence stage (Hearne, 2000; Northwest Regional Educational

Laboratory, 2003; Rhodes, Ochoa, & Ortiz, 2005; Roseberry-McKibbin, 2002). At this time, they have about 3000 words and use simple sentences to communicate with others. They also can have simple conversations with others and can ask elementary questions. As their speech production increases, however, grammatical errors persist. From about two to three years of second language exposure, children acquiring a second language fall into the intermediate language proficiency stage. At this time, these children have about 6000 words and have satisfactory face-to-face conversational proficiency. They are able to speak in more complex sentences, give their opinions, communicate their ideas, and do so with few grammatical errors. Children who are in the advanced language proficiency stage are able to utilize the second language in a manner commensurate with same-age native speakers.

Also of note, Cummins (1984) suggested that there are two categories of language proficiency that develop within different time spans: *basic interpersonal communication skills* (BICS) and *cognitive academic language proficiency* (CALP). BICS refers to the language used in social, informal contexts and can be acquired in two to three years. CALP comprises the language skills necessary to complete schoolwork and it takes five to seven years to develop.

Of further importance, a second language may be acquired at different points of time across childhood, resulting in some sequential differences (Owens, 2005). *Simultaneous acquisition* refers to the development of both languages before the age of three. This process of language acquisition begins with language mixing. Then, the child gradually makes the distinction between the languages and there is a growing awareness of their differences. In the final stages, there is a definitive division of the languages'

phonological and grammatical systems during which a dominant language may be evident through the child's use of vocabulary and idioms. *Successive acquisition* occurs more frequently and refers to the development of one language from birth and another after the age of three, usually in school. Development of the second language in this context begins with the child experiencing and utilizing the second language in social relationships. Here, the child's motivation is attempting to participate in a social group and he or she does so by emulating basic linguistic patterns he or she has observed. In the next stage of second language successive acquisition, communication is of chief importance to the child. He or she uses the basic units of the language comprehended in order to promote his or her communicative message. It is then in the final stage that children focus on producing the appropriate language form. When a second language is acquired in this fashion, there is the risk that early experience with the second language might impact the development of the first language. This can lead to a low proficiency in both languages. Conversely, it is also possible that first language comprehension can help lay the groundwork for the development of the second language due to the child's understanding of fundamental language processes. Cummins (1979) asserts that the development of a second language is dependent upon the level of competence in one's first language. Further, he suggests that in order to benefit from exposure to a second language and avert the potentially harmful affects of semilingualism, one must have reached a particular threshold of competence in one's first language. Therefore, a child must be sufficiently competent in his or her first language in order to satisfactorily acquire a second language without the potential for negatively affecting his or her skills in both languages.

Shames and Anderson (2002) refer to bilinguals as those who possess “various levels of proficiency in English, plus one...other [language]” (p. 596). Within this category, one may be a balanced bilingual, partial bilingual, or semilingual (Cummins, 1981; Owens, 2005). A balanced bilingual is equally proficient in both languages. Partial bilinguals are those who are more dominant in one language than the other. Semilinguals are those who are only semi proficient in both languages. Thus, it follows that Hispanic students are likely to progress through school as bilinguals, but with varying degrees of proficiency in both English and, likely, Spanish.

Language Demands in School

In order to function academically, socially, and emotionally in school, children must be able to adequately negotiate the complex language demands their environment presents. Shames and Anderson (2002) outline four key language-based skills that children must be proficient at in order to be successful in school: (1) representing abstract and remote events, (2) combining smaller components into larger ones, (3) processing different components simultaneously, and (4) moving across modalities and adapting to changes.

First, children must be able to understand how language is used to represent remote and abstract concepts (Shames & Anderson, 2002). This allows them to interpret causal relationships and learn new ideas. For example, students must comprehend how to carry out mathematical procedures, make connections between story characters’ emotions and actions, and learn about past events in their country’s history.

Next, children should be able to integrate the structural and functional units of language to produce meaningful communication. The structural part of language refers to

words and word parts, while function refers to the purpose of communicative attempts (Shames & Anderson, 2002). More fundamentally, students must be able to successfully combine basic structural units of language to create larger structural units and do the same for the functional language components. At the most basic level, students need to be able to form words and sentences (both orally and in written form) that convey meaning to those around them.

On the other hand, children must comprehend similar communicative initiatives from others (e.g., their teachers and peers). This requires that students process the different components of language (i.e., structural and functional units) concurrently in order to ascertain meaning from their communicative partners (Shames & Anderson, 2002). Furthermore, they must be able to relate incoming information to previously acquired information. For example, this is necessary for students to successfully understand stories and to engage in class-wide discussions.

Finally, children must be able to move between different communication modalities (e.g., writing, reading, listening, speaking) and link information acquired in one modality to that presented in others (Shames & Anderson, 2002). This is important for tasks such as writing responses to reading and discussing a story with peers. Also, children are required to react adaptively to shifts that arise during communicative dialogue. For example, a teacher may shift from a lecture to requiring students to work on a related task in small groups. This results in a shift in the language demands placed on the students that they must promptly adjust to. These four important skills must often be utilized concurrently, which requires children to be fluent and proficient in carrying out each of them.

Speech or Language Impairment (SLI)

Bedore and Pena (2008) define language impairment as “the inability to learn language as manifested by deficits in expressive and/or receptive language skills relative to age-matched peers who have comparable language exposure” (p. 1). Shames and Anderson (2002) also add that the impairment cannot be due to cognitive, motor, sensory, or socioemotional deficits. Expressive language refers to the ability to produce language, while receptive language is the ability to comprehend language. Children with expressive language issues may have difficulty with word retrieval, learning and applying grammatical rules, indicating causal relationships, producing “well-formed” sentences, and elaborating on ideas. They may have particular difficulty attempting to relay more complex information (Fey, 1986). Children with receptive language problems are unable to successfully process or assign meaning to the language of their communicative partner, whether it be presented in written or oral form. They may have trouble following instructions, acquiring information through a language medium, making connections in stories, processing ideas through language, and recognizing contradictions in passages (Shames & Anderson, 2002). Students with mixed expressive and receptive problems have trouble in both areas, though they tend to have more difficulty using language in a communicative way than comprehending language.

In Section 300.8(c) of The Individuals with Disabilities Education Improvement Act (IDEIA, 2004) a Speech or Language Impairment (SLI) is defined as a “communication disorder, such as stuttering, impaired articulation, a language impairment, or a voice impairment that adversely affects a child’s educational performance.” Furthermore, deficits cannot be due to differences in culture, limited

English proficiency, or limitations in educational experience. The remainder of this document will use the IDEIA's more broad classification terminology (i.e., SLI), but will concentrate specifically on the issues surrounding the assessment and classification of language impairment as it relates to Hispanic children in schools. This focus stems from the difficulties that have arisen in distinguishing language differences from language disorders in culturally diverse children, including Hispanic students.

Children with SLI often experience difficulties in school that result from their partial or total inability to master one or more of the aforementioned skills: those outlined by Shames and Anderson (2002). Language is the vehicle from which information is transferred and learned in both oral and written form. Evidence of a SLI in school may include difficulty with learning new information, reading and writing, and/or social relationships, all of which are closely related to language (Bashir, Conte, & Heerde, 1998; Shames & Anderson, 2002). Thus, students with SLI may have difficulty learning new material and expressing their own ideas in one or both of these forms, including in social contexts. This is especially true as concepts increase in complexity.

Assessment Procedures for SLI

Because language impairment may involve deficits in receptive language, expressive language, or both, SLI assessment procedures should evaluate both of these domains (Kummerer, Lopez-Reyna, & Hughes, 2007; Tyler & Tolbert, 2002; Washington & Craig, 2004). Some states require tests of both receptive and expressive language when evaluating SLI (e.g., NJ Special Education Code, 2006), while others require using either or both depending on relevant referral information (e.g., Connecticut State Department of Education, 2008; Wisconsin Department of Public Instruction,

2003). Such tests may measure solely receptive or expressive language, or may be more comprehensive measures of both. It follows, then, that the subsequent discussion of standardized tests of language refers to receptive, expressive, and comprehensive (i.e., receptive and expressive) measures of language.

SLI is typically assessed based on the assertion that children with language impairment will score lower on measures of language than do typically-developing children (Gutierrez-Clellen & Pena, 2001; Spaulding, Plante, & Farinella, 2006). As a result, professionals frequently utilize cutoff scores on standardized tests of language to guide diagnosis. That is, children are often diagnosed with language impairment if they score below a certain criterion and, thus are shown to have substantially more impairment than others their age. Perhaps in part because the Individuals with Disabilities Education Improvement Act (2004) requires that all evaluation instruments be technically sound, public school systems appear to be particularly dependent on the use of standardized language tests to classify children with SLI (Oetting, Cleveland, & Cope, 2008; Spaulding, Plante, & Farinella, 2006). In particular, schools often set cutoff scores to determine which students are eligible for special education under SLI, such as scoring -1, -1.5, -1.75, or -2 standard deviations below the mean on a standardized test of language (Oetting, Cleveland, & Cope, 2008). It must be noted that IDEA guidelines state that single assessment measures should not be used as the sole indicator of disability status. However, in this review, the emphasis is placed on the use of standardized tests of language because of the seemingly substantial role they play in the classification of students with SLI in schools.

Though this method of assessment seems to be relied on in schools, these traditional procedures are questionable, especially when evaluating culturally diverse children for SLI. First, the use of cutoff scores to ascertain the presence or absence of language impairment has been shown to be problematic, especially when such criteria are not empirically-derived (Merrell & Plante, 1997; Oetting, Cleveland, & Cope, 2008; Spaulding, Plante, & Farinella, 2006). Spaulding, Plante, and Farinella reviewed 43 commercially available, standardized tests of language in order to determine if the use of cutoff scores was an appropriate diagnostic tool. Utilizing the test manuals, the authors compared the scores of children with language impairment to matched typically-developing children within the normative samples. Based on their analyses, the authors found that the degree to which children with language impairment scored lower than typically-developing children varied based on the individual tests. Also, Oetting, Cleveland, and Cope found that using arbitrary cutoff scores to evaluate SLI in culturally diverse students was similarly ineffective, while using empirically derived criteria might be more useful. Furthermore, no standardized assessment measures utilize norm samples that control for cultural experience (e.g., acculturation) or bilingualism (Rhodes, Ochoa, & Ortiz, 2005). When using such instruments, culturally diverse students' test scores are derived based on a norm sample with a dissimilar experiential background (Salvia & Ysseldyke, 1991). This indicates that comparing their scores to a cutoff score derived from this norm sample is futile. These factors support the premise that assessing SLI utilizing arbitrary cutoff scores (e.g., one standard deviation below the mean), as is frequently done in school systems, does not result in accurate classifications of SLI. In fact, one might surmise that the inconsistency in cutoff scores used across different

schools and districts is indicative of their “arbitrary nature” (Spaulding, Plante, & Farinella, 2006, p. 61).

These concerns are especially salient when examining SLI assessment procedures for Hispanic students; differences in culture may make it difficult to discern language difference from language disorder in diverse students, especially when using standardized instruments (Gillam, Pena, & Miller, 1999; Marbley, Bonner, & Berg, 2008; Miller, Gillam, & Pena, 2001; Shames & Anderson, 2002). Miller, Gillam and Pena define a language difference as “normal language development that appears unusual due to social, cultural, and experiential factors” (p. 3). This phenomenon is different from a language disorder in which language development is atypical due to problems with language learning processes. To start, it has been shown that Hispanic students perform worse on language assessment procedures that are less consistent with their cultural language styles (Pena & Quinn, 1997; Pena, Quinn, & Iglesias, 1992). On such standardized measures, children from nonmainstream cultures (e.g., Hispanic students) may score lower than is expected due to limited English proficiency, differences in language experiences, insufficient mainstream language learning opportunities, or differences in language styles (Artiles & Trent, 1994; Gutierrez-Clellen & Pena, 2001; Harry, 1994). Findings suggest a pattern of language differences rather than true language disorders in culturally diverse children. For example, Heath (1982, 1986) found that Mexican-American immigrants did not directly teach their children some routine, language-based interactions that are commonly utilized in mainstream schooling; such mainstream language conventions include asking children to respond to questions regarding known information and reviewing known information (Anderson-Yockel & Haynes, 1994; Pena & Quinn, 1997).

Rather, Mexican-American parents were more likely to ask questions in order to obtain unknown information. Also, perhaps due to such language modeling by their parents, Hispanic children have been found more likely to utilize functions to label objects, rather than using specific objects names (Gutierrez-Clellen & Iglesias, 1987, as cited in Pena, Iglesias, & Lidz, 2001; Heath, 1982; Pena, Quinn, & Iglesias, 1992). As a result of such language differences, Hispanic children may not perform as well as children from mainstream cultures on standardized language tests, which inevitably require test-takers to respond to items indicating both word labels and previously-learned, known information (Langdon & Chang, 1992). Less research has been conducted on the opposite phenomenon: when a language disorder does exist, but, due to insufficient assessment procedures, a language difference is assumed. Though less seems to be known in this regard, Gutierrez-Clellen and Pena (2001) warn that this occurrence might result in inaccessibility to the appropriate services. That is, when children with language disorders are not classified with SLI, they might be unable to receive the appropriate services they would be eligible for under that classification. Overall, it is evident that utilizing traditional assessment instruments as the key indicators of SLI may be problematic because these measures appear to lack the sensitivity necessary to differentiate language difference from language disorder. Gillam, Pena, and Miller (1999) caution that even pairing traditional, standardized assessment measures with real-time observation may not produce the necessary sensitivity. Because the evaluation of SLI involves culturally loaded tests with high language demands, Hispanic children are especially at risk for their language abilities to be evaluated inappropriately (Harry, 1994).

These concerns underscore the need for more culturally sensitive assessment procedures, especially when examining language ability in Hispanic children. This is especially critical in order to prevent the negative outcomes associated with misclassification. Furthermore, schools are legally required to choose and administer assessment procedures in a manner that is not discriminatory on the grounds of race or cultural background (IDEA, 2004). In regard to SLI, the most critical need appears to be for a procedure that allows professionals to tease apart language differences from language disorders in Hispanic children (Gillam, Pena, & Miller, 1999; Gutierrez-Clellen & Pena, 2001; Marbley, Bonner, & Berg, 2008; Miller, Gillam, & Pena, 2001; Salend, Duhany, & Montgomery, 2002). Gillam, Pena, and Miller assert that doing so requires procedures designed to evaluate the child's ability to learn language and to respond to language interventions. Including these features seems critical when assessing SLI, especially with Hispanic students. It is only with this accurate evaluation of student strengths and needs that appropriate programming and, thus, academic growth will follow.

Dynamic Assessment

Dynamic Assessment is one approach to assessment that has been identified as promising in the development of more culturally competent evaluations for SLI (Gutierrez-Clellen & Pena, 2001; Haywood & Lidz, 2007; Lidz, 1991; Pena, Iglesias, & Lidz, 2001; Pena et al., 2006). Vygotsky's (1978, 1986) zone of proximal development (ZPD), part of his total cognitive developmental model, became the foundation upon which dynamic assessment was developed. The ZPD refers to the "distance between the actual developmental level as determined by independent problem solving and the level

of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978, p. 86). Knowing an individual’s ZPD in a particular area, then, is thought to be a helpful indicator of the child’s learning potential. More specifically, Vygotsky’s model posited that psychological processes are influenced by an individual’s interaction with adults and more skilled peers, particularly when they are teaching the individual a new concept. Thus, dynamic assessment has come to focus on generating a formulation of a child’s ZPD in a certain domain in order to gain a richer determination of his or her strengths and needs (Miller, Gillam, & Pena, 2001). Further, Pena et al. (2006) propose that the Vygotskian style of dynamic assessment would include measures of performance before and after adult instruction. Capable learners should produce better results on an activity after adult intervention due to the intervention’s influence on the learners’ psychological processes. The implication here is that those who are not competent learners (e.g., children with SLI) would be unable to progress or would progress minimally on such a task despite instruction from a more competent partner.

In modern applications of the model, dynamic assessment refers to an assessment approach in which the evaluator actively intervenes in order to gauge the individual’s responsiveness to intervention and learning potential (Haywood & Lidz, 2007). Haywood and Lidz outline five core principles that are essential to the dynamic assessment model:

1. Some abilities that are important for learning (in particular) are not assessed by normative, standardized intelligence tests.
2. Observing new learning is more useful than cataloguing (presumed) products of old learning. History is necessary but not sufficient.

3. Teaching within the test provides a useful way of assessing potential as opposed to performance.
4. All people typically function at less than their intellectual capacity.
5. Many conditions that do not reflect intellectual potential can and do interfere with expression of one's intelligence¹. (p. 7)

As such, dynamic assessment offers a representation not of the child's abilities as compared to their peers, but of his or her own abilities compared across time, and with the influence of teaching. One important characteristic of dynamic assessment is its ability to gather information that is less affected than standardized assessment by environmental factors, including cultural differences or educational opportunity; it is thought to give an indication of learning potential rather than a static measure of current performance. The assessment information that yields this insight from a dynamic assessment procedure includes measures of current level of performance, the nature of the intervention needed to generate improvement (i.e., how much? what kinds?), and data on how the child responded to the intervention (Haywood & Lidz, 2007). The difference between the current level of performance and the level of performance following the intervention reflects the child's ZPD, while the nature of the intervention necessary to affect change reflects the child's modifiability within that domain. Combining these two factors appears to be especially useful in making recommendations and in making disability classification, especially in regard to SLI (Haywood & Lidz, 2007; Lidz, 1991;

¹ It is important to recognize that these concepts refer specifically to the psychological domain, but that Haywood and Lidz reference the use of DA within other domains, including speech and language.

Miller, Gillam, & Pena, 2001; Pena, Iglesias, & Lidz, 2001; Pena, Quinn, & Iglesias, 1992).

One prominent application of the dynamic assessment approach is the test-teach-retest model (e.g., Budoff, 1987; Feuerstein, Rand, Jensen, Kaniel, & Tzuriel, 1987; Lidz, 1991; Miller, Gillam, & Pena, 2001; Pena et al., 2006). The dynamic assessment procedure in this case starts with the implementation of a pretest in order to gather information on the individual's baseline level of functioning in the relevant domain. The examiner then implements an intervention during one or two sessions, after which the child's modifiability is rated. Feuerstein (1979) first referred to these teaching sessions as Mediated Learning Experiences (MLEs) and defined them as follows:

...the interactional processes between the developing human organism and an experienced, intentioned adult who, by interposing himself between the child and external sources of stimulation, "mediates" the world to the child by framing, selecting, focusing, and feeding back environmental experiences in such a way as to produce in him appropriate learning sets and habits. (p. 71)

Thus, the MLEs are designed to guide the child with the goal of producing improvement within the relevant domain. Each MLE must comprise at least five mediation characteristics: intention to teach, meaning, transcendence, planning, and transfer (Feuerstein, 1980; Lidz, 1991; Miller, Gillam, & Pena, 2001). Intention to teach strategies refer to the examiner's conscious determination to impact the child and how this is communicated to him or her. This may be manifested in the assessor's attempts to orient the child toward the goals of the activity. Meaning strategies are utilized in order to highlight for the child the areas where his or her attention should be concentrated, and to

assign value and interpretation to relevant pieces of the task. Next, transcendence strategies provide task relevance to the child past that of the immediate environment. The child may be encouraged to consider “what if” or “cause-effect” scenarios in order to accomplish this. Then, the examiner uses planning strategies to encourage the child to consider what he or she has learned and what steps he or she can take to successfully apply this knowledge in future, similar situations. Finally, transfer strategies refer to the examiner’s summarization of the MLE at the conclusion of the session. This is done in order to encourage the child to consider utilizing the newly learned concepts in the future. Once the evaluator engages the individual in one or two MLE sessions, the child is given a modifiability rating that quantifies the effort required to evoke improvement and the child’s responsiveness within the MLEs. A posttest is then administered to measure the child’s performance following the intervention and, more specifically, to examine how much change resulted from the mediation.

Because the dynamic assessment model asserts that competent learners are those who are able to learn within the process, this approach has implications for disability classification. Miller, Gillam, and Pena (2001) state that an individual can be considered a competent learner if they are rated at least as moderate on modifiability following MLE sessions, and if they made significant improvements from pre to post testing. Improvements are said to be significant if the child’s ratings increase in more than one area, including both those targeted and not targeted during the MLE sessions. A child would also be considered competent if he or she obtains very high scores at pretest. Conversely, the individual may have learning difficulties indicative of disability if he or she receives low modifiability scores and does not improve significantly from pre-to-post

testing (i.e., his or her ratings do not improve following MLEs or they only improved slightly: small improvements and only in the areas specifically targeted). These kinds of guidelines have resulted in the utilization of dynamic assessment as a method of differentiating between students with and without SLI. Dynamic assessment appears to be an especially important tool for distinguishing language differences from language disorders in culturally diverse children.

In this context, dynamic assessment is typically implemented using the test-teach-retest formula described previously (Guitierrez-Clellen & Pena, 2001). For example, Pena, Quinn, and Iglesias (1992) used a dynamic assessment procedure to evaluate the word learning abilities of African American and Puerto Rican preschoolers. The authors found that children with and without possible language disorders scored similarly on traditional measures of assessment (i.e., standardized tests). However, posttest scores following mediation procedures were a great deal higher for those without possible language disorders. Students with possible language disorders showed fewer gains from pre to post tests and needed greater adult effort to produce change during mediation sessions (i.e., they had lower modifiability ratings). Pena, Iglesias, and Lidz (2001) then attempted to expand upon this work by adding a no-mediation control group to their study. In their study, the children participating in the MLE condition improved significantly more from pre to posttests than those not exposed to the MLE condition. They also found that children with typical language development demonstrated significantly greater improvements than children with low language ability. Moreover, the authors found that using solely pretest scores to classify students with low language ability resulted in “unacceptably high” false positive rates. They warned that children

from culturally diverse backgrounds might be particularly susceptible to this misclassification when using traditional, static tests to evaluate language ability. In contrast, a combination of posttest scores and modifiability ratings resulted in most accurate classification of language ability. Lidz and Pena (1996) and Guitierrez-Clellen and Pena (2001) also presented case examples of similar dynamic assessment procedures and their apparent utility in differentiating children with and without language disorders.

Finally, Pena et al. (2006) used dynamic assessment of narrative abilities with culturally diverse first and second grade students. They suggested that narrative assessment had a high content validity because children are frequently required to use and comprehend narratives as home and at school. Moreover, narratives have been utilized as a measure of oral language ability, including when diagnosing language impairment (Kaderavek & Sulzby, 2000; McCabe, Bliss, Barra, & Bennett, 2008; Tyler & Tolbert, 2002). Pena et al. found that children's performance on narratives improved following MLEs. Children with and without language disorders differed with respect to both modifiability ratings and improvement from pre-to-post testing. Children without language disorders made greater gains and were given higher modifiability scores. Furthermore, Pena et al. were unable to accurately classify children with language disorders using pretest scores; doing so in various permutations resulted in either "unacceptably high" false-positive or false-negative rates. In particular, the African American and Hispanic children in the study were much more likely than European American children to receive low scores on pretest measures. The authors found that using a combination of modifiability ratings and posttest scores produced much more accurate classification results. Furthermore, the best single predictor of language ability

was modifiability ratings. They concluded that their study, along with other similar studies, indicated that dynamic assessment of language ability might be a valuable way to evaluate culturally diverse students for SLI. Traditional assessment measures are likely to result in mistakenly deflated scores for this population due to the environmental factors discussed previously.

The purpose of the current study is to compare a traditional standardized measure of language and a dynamic assessment procedure for assessing language ability of Hispanic, bilingual first grade students. First, the study will differentiate the students as typically-developing or at risk for SLI classification based upon a standardized assessment measure, as is often done in schools. Then, it will examine pre to posttest gains made in a dynamic assessment procedure by children previously deemed typically-developing and at risk for SLI classification using traditional means. The study will also compare the modifiability ratings of typically-developing versus at-risk students. In doing so, this study will explore the potential place for dynamic assessment within the SLI assessment process, especially as a way to produce a more culturally sensitive evaluation for CLD children.

CHAPTER III

METHOD

Participants

The participants in this study were a subsample of 475 Hispanic children taking part in a three-year longitudinal study that investigated social-emotional and academic resilience of young children. A total of 25 children attending first grade in an urban school district in the Northeast participated in the current study. The students ranged in age from 6 to 7 ($M = 7.04$, $SD = .31$) and 48% were male. Most of the students participating in this study were randomly selected. Past test scores were screened in order to select students more likely to meet criteria for the at-risk group. All participants were Hispanic-American bilinguals and qualified for free or reduced fee lunch.

Measures

Oral Language: All children were given the Woodcock-Munoz Language Survey-Revised (WMLS-R) in English (Woodcock, Munoz-Sandoval, Ruef, & Alvarado, 2005). The WMLS-R is a broad measure of proficiency in oral language, language comprehension, reading, and writing (Alvarado, Ruef, & Schrank, 2005). It is marketed as a tool for assessing English and Spanish language proficiency in students for whom English is a second language. It is made up of seven subtests that combine in different permutations to generate eleven composite scores. For each composite, a CALP score is determined. The Oral Language—Total composite of the WMLS-R is comprised of four

subtests (Picture Vocabulary, Verbal Analogies, Understanding Directions, and Story Recall) and includes both measures of receptive and expressive language. The median reliability for this composite is .95 with individuals age 5 to 19. The Broad English Ability—Total composite is comprised of all seven subtests: Picture Vocabulary, Verbal Analogies, Letter-Word Identification, Dictation, Understanding Directions, Story Recall, and Passage Comprehension. The median reliability for this cluster is .98 for individuals age 6 to 19.

Modifiability: Modifiability was rated based on the criteria outlined by Miller, Gillam, and Pena (2001). Student modifiability ratings were comprised of teaching effort (rated from 1-“a lot”- to 5-“little”) and student responsiveness (rated from 1-“not very”- to 5-“very”). The teaching effort and student responsiveness scores were added together to obtain a total modifiability score ranging from 2 to 10. Total modifiability scores were considered to be low (scores ranging from 2 to 5), moderate (scores ranging from 6 to 7), or high (scores ranging from 8 to 10).

Procedure

The Oral Language—Total of the WMLS-R was used as an indicator of language ability. CALP scores were derived from the Broad English Ability—Total as an indicator of English language proficiency. Students scoring more than one standard deviation below the mean on the Oral Language Total (i.e., below 85) were labeled as at risk for SLI classification, while those scoring 85 or higher were deemed typically-developing. A cutoff score was used to mirror the procedure frequently employed by school districts to classify SLI; particularly, one standard deviation below the mean of a standardized

measure of language is one cutoff score used in this setting (Oetting, Cleveland, & Cope, 2008; Spaulding, Plante, & Farinella, 2006).

All children then underwent a dynamic assessment of language ability utilizing the procedure outlined in Pena et al.'s (2006) study. They participated in a pretest, two 30-minute MLE sessions, and a posttest. Contrary to Pena et al.'s study, all four of these sessions were conducted within a one-and-a-half-week time period (each on a different day). In the pretest phase, children were presented with a wordless picture book, *Two Friends* (Miller, 1999b), and told, "This is my book, *Two Friends*. This book is a little different than other books you've seen because it has only pictures and no words. I want you to tell me a story that goes along with the pictures. I'm going to give you some time to look through the pictures before you tell me your story. Let me know when you're ready." The examiner recorded the children's stories by typing their responses onto a laptop computer as they spoke.

The children then each participated in two MLE sessions. The examiner used the scripts developed by Pena et al. (2006) that contained each of the five mediation characteristics: intention to teach, meaning, transcendence, planning, and transfer (Feuerstein, 1980; Lidz, 1991; Miller, Gillam, & Pena, 2001). The use of these scripts within Pena et al.'s procedural framework allowed for the presentation of the MLEs to be standardized, but permitted some flexibility in order for the examiner to address the unique demands of each child. Pena et al. designed the MLE sessions with the aim of increasing the length and complexity of the children's stories by targeting improvements in story components (i.e., setting, character information, temporal order of events, and causal relationships) and episode elements/structure (i.e., initiating event, internal

response, attempt, plan, consequence, and reaction/ending). Thus, nontargeted areas included story ideas and language, and listener effort. The first MLE was conducted in the context of the child's prettest story. That is, the examiner and child went over and worked through the story the child generated corresponding to *Two Friends*. During the second MLE, the examiner guided the child through developing a story for *One Frog Too Many* (Mayer & Mayer, 1975), another wordless picture book. Following the second MLE, the examiner rated the child on modifiability. Puppets or other toys were not used during MLE sessions, as in Pena et al.'s study.

At posttest, the same procedure as the prettest was used, but instead with the book, *Bird and His Ring* (Miller, 1999a). The books used for the prettest and posttest were presumed to measure language ability in an equivalent manner, as was demonstrated by Pena et al. (2006). In that study, the authors demonstrated that, without intervention, children told qualitatively and quantitatively similar stories when presented with the two different books: *Two Friends* and *Bird and His Ring*.

Scoring Procedures

Following posttest, the examiner scored the transcriptions based on the procedures outlined by Miller, Gillam, and Pena (2001); this protocol was also used in Pena et al.'s (2006) study. The narratives were rated based on the following categories: story components (i.e., setting, character information, temporal order of events, and causal relationships), story ideas and language (i.e., complexity of ideas, complexity of vocabulary, grammatical complexity, knowledge of dialogue, and creativity), episode elements and structure (i.e., initiating event, internal response, attempt, plan, consequence, and reaction/ending), and listener effort. Each element comprising the story

components and story ideas/language sections was rated on a 5-point Likert scale then summed together to obtain a total score for each. The episode elements/structure was rated on a 7-point scale, depending upon how many of those components the child included in his or her story. Listener effort was rated on a 5-point Likert scale. Each category was then summed to produce an overall, total story score ranging from 11 to 57 points. Measures of productivity were then calculated including the number of words, number of clauses, number of C-units (i.e., the main clause and its modifiers), mean length of C-units (MLC-unit; i.e., total number of words divided by the total number of C-units), and the number of clauses per C-unit.

Data Analysis Procedures

First, the mean Oral Language—Total scores were compared between the at-risk and typically-developing groups using an independent samples t-test. An independent samples t-test was also used to compare the mean CALP scores between the groups in order to determine whether or not there existed English language proficiency differences between the students.

To address the first research question, the mean pretest total story scores and productivity measures were compared to respective posttest measures for both the at-risk and typically-developing groups using a dependent samples t-test.

To address the second research question, an independent sample t-test was conducted in order to compare the mean gain scores between the two groups. An independent samples t-test was also performed to compare the mean modifiability ratings between both groups.

To address the third research question, Miller, Gillam, and Pena's (2001) procedure for distinguishing language abilities was applied to each individual in the study. According to these authors, an individual is designated a competent learner (i.e., typically-developing) if he or she is rated at least moderately on modifiability (i.e., total score between 6 and 10) and if he or she makes significant improvements following MLE sessions. Improvements are said to be significant if the child's ratings increase in more than one area (i.e., increase by at least one level/point), including both those targeted and not targeted during the MLE sessions. A child would also be considered competent if he or she obtains very high scores at pretest. The individual may have learning difficulties indicative of a disability (i.e., at risk for SLI classification) if he or she receives low modifiability scores (i.e., total score between 2 and 5) and does not improve significantly from pre-to-post testing (i.e., his or her ratings do not improve following MLEs or they only improve slightly: one level/point only in the areas specifically targeted). Case examples presented by Miller, Gillam, and Pena were also utilized in order to help make such judgments when necessary. The manner in which these results indicate who is and is not at risk for SLI classification was then compared to that of the WMLS-R. Judgments were then made regarding implications for the SLI assessment in Hispanic students.

CHAPTER IV

RESULTS

Zero Order Correlations

To understand the relationships among the variables of interest, zero order correlations are displayed in Table 1. WMLS-R Oral Language—Total scores were not significantly correlated with any of the dynamic assessment measures. Modifiability was highly correlated with all other dynamic assessment measures, indicating that the higher the modifiability score, the higher the scores in the other dynamic assessment variables. These variables include the total story score and all productivity measures (i.e., number of words, number of clauses, mean length of C-unit, number of C-units, and number of clauses per C-unit). The measures of productivity were also significantly correlated with each other.

Oral Language and English Proficiency

Means and standard deviations of oral language, English proficiency, and modifiability measures are displayed in Table 2. Because the sample sizes for the at-risk and typically-developing groups differed, Levene's Test for Equality of Variances was performed. Levene's test was not statistically significant for the measures of language ability (i.e., Oral Language—Total scores) or English proficiency (i.e., CALP scores). The results of the independent samples t-test indicate that the mean oral language scores for the at-risk group were significantly different than those of the typically-developing

group ($t(23) = 6.18, p < .001$). Specifically, the students assigned to the at-risk group scored significantly lower than those in the typically-developing group. The magnitude of the difference in means was large ($\eta^2 = .62$). Further, students deemed at-risk for SLI classification had significantly lower English proficiency scores than those judged to be typically-developing ($t(23) = 3.85, p = .001$). The magnitude of the difference in means was large ($\eta^2 = .39$).

Table 1
Zero Order Correlations Among Variables of Interest

Variable	1	2	3	4	5	6	7	8
1. Oral Language—Total	1.00							
2. Modifiability	.31	1.00						
3. Total Story Score	.35	.56**	1.00					
Productivity								
4. Number of Words	.13	.52**	.86**	1.00				
5. Number of Clauses	.15	.53**	.82**	.96**	1.00			
6. Mean Length C-unit	.28	.41*	.78**	.69**	.54**	1.00		
7. Number Clauses per C-unit	.30	.42*	.68**	.53**	.55**	.80**	1.00	
8. Number of C-units	-.00	.42*	.62**	.86**	.90**	.23	.13	1.00

** $p < .01$, * $p < .05$

Table 2
Descriptive Statistics for Typically-Developing Children and Children At Risk for SLI Classification

Variable	At-Risk <i>n</i> = 7		Typically-Developing <i>n</i> = 18	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Oral Language—Total	78.71	4.27	99.72	8.51
CALP Level—English	3.71	.39	4.53	.50

Note: The Oral Language—Total Composite of the WMLS-R has a *M* = 100 and a *SD* = 15. CALP levels are as follows: 1 (Negligible), 2 (Very Limited), 3 (Limited), 4 (Fluent), 5 (Advanced), 6 (Very Advanced). Modifiability ratings are as follows: low (2-5), moderate (6-7), high (8-10).

Research Question 1

The study sought to determine if the students made significant improvements in measures of their language abilities during a dynamic assessment procedure. Means and standard deviations of the total story scores and productivity measures at pretest and posttest for each group are displayed in Table 3. Productivity measures include: number of words, number of clauses, number of C-units, mean length of C-units, and number of clauses per C-unit. In order to describe the individual variability in productivity measures, especially in number of words produced, Figure 1 portrays individual pretest and posttest scores for the at-risk group and Figure 2 depicts the same for the typically-developing group. It should be noted that one student in the at-risk group produced a great deal more words in her stories, both at pretest and at posttest, than the others in this group. However, this student appeared to make similar gains in this area from pretest to posttest, and scored similarly to the others in measures of oral language.

For the at-risk group, the results of the paired-samples t-test suggest that the mean posttest total story score was significantly higher than the mean pretest total story score

($t(6) = 12.23, p < .001$). The magnitude of this difference was large ($\eta^2 = .96$). The mean number of words used was significantly higher at posttest than at pretest ($t(6) = 4.49, p = .004$). The magnitude of this difference was large ($\eta^2 = .77$). Students produced, on average, a significantly higher number of clauses at posttest than at pretest ($t(6) = 3.16, p = .02$). The magnitude of this difference in means was large ($\eta^2 = .62$). Their mean length of C-unit was also significantly higher at posttest than at pretest ($t(6) = 4.15, p = .006$). The magnitude of this difference was large ($\eta^2 = .74$). There were no significant differences in means between pretest and posttest in the number of clauses produced per C-unit and the number of C-units produced.

For the typically-developing group, the results of the paired-samples t-test suggest that the mean posttest total story score was significantly higher than the mean pretest total story score ($t(17) = 8.55, p < .001$). The magnitude of this difference of means was large ($\eta^2 = .81$). The mean number of words used was significantly higher at posttest than at pretest ($t(17) = 3.54, p = .002$). The magnitude of this difference was large ($\eta^2 = .42$). Students produced, on average, a significantly higher number of clauses at posttest than at pretest ($t(17) = 2.68, p = .016$). The magnitude of this difference was large ($\eta^2 = .30$). Their mean length of C-unit was also significantly higher at posttest than at pretest ($t(17) = 2.41, p = .028$). The magnitude of this difference in means was large ($\eta^2 = .25$). Students produced, on average, significantly more C-units at posttest than at pretest ($t(17) = 2.45, p = .025$). The magnitude of this difference was large ($\eta^2 = .26$). There were no significant differences in means between pretest and posttest in number of clauses produced per C-unit.

Research Question 2

Next, the performance of the at-risk and typically-developing groups were compared to one another. It was of particular interest to examine whether their improvement from pretest to posttest differed and whether they performed differently on ratings of modifiability. Means and standard deviations of gain scores for each variable are presented in Table 4. Descriptive statistics of modifiability ratings are given in Table 5. Levene's Test for Equal Variances indicated that the variances in modifiability for the at-risk and typically-developing groups are not significantly different from one another. The results of the independent samples t-test suggest that there are no significant differences in mean modifiability ratings between the at-risk and typically-developing groups. Specifically, both groups were rated, on average, at the high end of the moderately modifiable range.

Levene's Test for Equal Variances was significant for gain scores and, thus, unequal variances were assumed. The independent samples t-test performed indicates that there was no significant difference in mean pretest to posttest total story gain scores between the at-risk and typically-developing groups. Productivity measures were then examined. Unequal variances were also assumed for the mean gains in number of words for the at-risk and typically-developing groups, due to a significant result in Levene's test. The results of the independent samples t-test suggests that the mean gains in number of words from pretest to posttest were not significantly different between the at-risk and the typically-developing groups. Levene's test was not significant for the remaining productivity measures; equal variances were assumed. The results of the independent samples t-test indicate that there were no significant differences between the at-risk and

typically-developing groups when examining mean pretest to posttest gain scores in the productivity measures: number of clauses, mean length of C-unit, number of C-units, and number of clauses per C-unit.

Table 3
Descriptive Statistics for Total Story and Productivity at Pretest and Posttest for
Typically-Developing Children and Children At risk for SLI Classification

Variable	At-Risk <i>n</i> = 7				Typically-Developing <i>n</i> = 18			
	Pretest		Posttest		Pretest		Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Total Story Score	20.00	7.94	31.14	8.88	25.61	7.59	34.22	8.40
Story Components	6.14	3.18	10.20	3.68	8.56	3.13	11.50	4.09
Story Ideas and Language	9.43	4.04	13.86	4.30	10.94	3.02	15.00	3.09
Episode Elements and Structure	3.00	1.29	4.00	1.00	3.44	1.29	4.44	1.20
Listener Effort	1.43	.53	3.00	1.00	2.67	1.24	3.28	1.02
Productivity								
Number of Words	114.14	94.18	163.86	88.39	126.17	47.25	171.89	67.85
Number of Clauses	21.57	15.11	28.86	15.25	23.22	7.43	30.00	12.59
Mean Length of C-unit	6.19	1.40	7.84	1.13	6.88	2.00	7.77	1.30
Number C units	17.43	11.03	20.14	7.95	18.17	3.99	22.00	7.81
Number of Clauses per C-unit	1.22	.20	1.39	.26	1.27	.26	1.35	.20

Note: Total Story Score is comprised of the following: Story Components, Story Ideas and Language, Episode Elements and Structure, and Listener Effort.

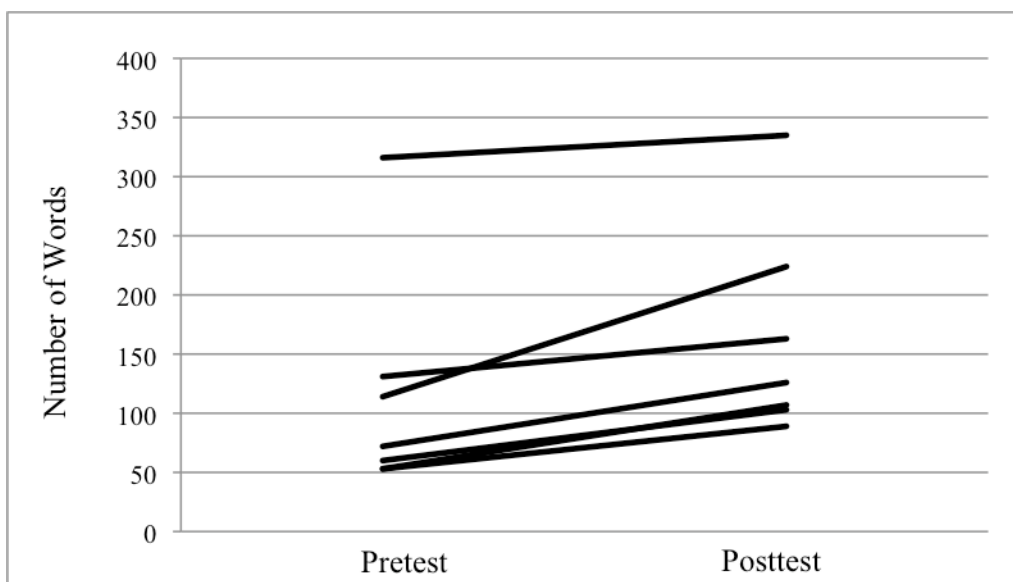


Figure 1. Growth in Number of Words for Children At Risk for SLI Classification.

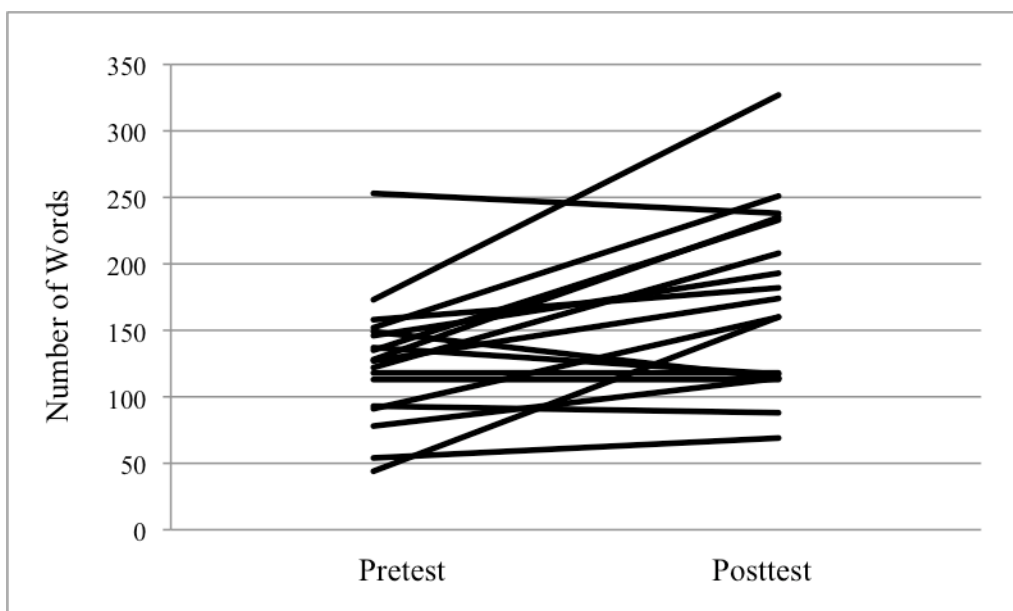


Figure 2. Growth in Number of Words for Typically-Developing Children.

Table 4
Descriptive Statistics of Gain Scores for Typically-Developing Children and Children At Risk for SLI Classification

Variable	At-Risk <i>n</i> = 7		Typically-Developing <i>n</i> = 18	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Total Story Score	11.14	2.41	8.61	4.27
Productivity				
Number of Words	49.71	29.33	45.72	54.74
Number of Clauses	7.29	6.10	6.78	10.71
Mean Length of C-unit	1.64	1.05	.89	1.57
Number C-units	2.71	5.53	3.83	6.63
Number of Clauses per C-unit	.17	.30	.07	.20

Table 5
Descriptive Statistics of Modifiability Ratings for Typically-Developing Children and Children At Risk for SLI Classification

Variable	At-Risk <i>n</i> = 7		Typically-Developing <i>n</i> = 18	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Modifiability	7.71	1.50	7.94	1.89

Research Question 3

Finally, this study examined the worth of dynamic assessment of language ability in evaluating Speech or Language Impairment in culturally and linguistically diverse students. To address the third research question, Miller, Gillam, and Pena's (2001) procedure for determining language ability was utilized at the individual student level. That is, this criteria for discerning competent language learners (i.e., typically developing students) from those with language learning difficulties (i.e., students at risk for SLI

classification) was applied to each student in the study. When using this dynamic assessment procedure to make judgments about language ability, none of the students in the at-risk group were assessed to have language-learning difficulties. On the contrary, all students were judged competent language learners; they were rated at least moderate on modifiability and made significant improvements following MLEs, or they had very high scores at pretest. Improvements are said to be significant if the child's ratings increase in more than one area (i.e., increase by at least one level/point), including both those targeted and not targeted during the MLE sessions. Scores are said to be very high if the child's ratings are above moderate (i.e., above the midpoint level) on most or all of the components.

These improvements are evident when examining the students' individual responses. For example, while one student's story at pretest read more like a list of events, his posttest story was more cohesive, had more detail, and was easier for the reader to follow. In both wordless picture books, the main character encounters other characters on their way to solve a problem. At pretest, this student described these events, in part, as follows: "And the dog found a dragon. And the then the dog saw a animal." Conversely, at posttest he gave, in part, the following description: "Tom said to a porcupine, 'You saw a ring around here?' Then the porcupine said no and then the iguana was in the ground and moving his head. And then Tom said to the alien, 'You saw a ring around here?'" While the latter response contains some superfluous information, it also provides a clearer explanation than the pretest story of what the main character is doing when he comes across the other characters and why. These are indicators of improvements in both targeted (episode elements and structure) and non-targeted areas

(story ideas and language, and listener effort). Further, his pretest story began, “The dog is saying hi to the cat,” which is merely a description of the first picture in the book and does little to make his story more complete. Conversely, the posttest story began, “One morning, Tom saw a ring,” a statement that indicates setting, background information related to future story details, and a character detail. Such components signify improvements in both targeted (story components, episode elements and structure) and non-targeted (story ideas and language, listener effort) areas.

Within the typically developing group, 2 of 18 students, or 11.11% were deemed at-risk for SLI classification (i.e., they received low modifiability scores and did not make significant improvements following MLEs). It was determined that the results for one student in this group were inconclusive and in need of further assessment, as they did not follow any clear pattern outlined by Miller, Gillam, and Pena. Specifically, this student received a low modifiability score, but showed both improvement and regression in targeted, as well as non-targeted areas. Thus, 15 of the 18 students in the typically-developing group, or 83.33%, were determined to have typical language abilities.

Both improvement and lack thereof is apparent when examining the respective student stories. For example, one student in the typically-developing group judged to have language learning difficulties by dynamic assessment did not demonstrate improvements in targeted areas. In particular, in both her pretest and posttest stories, there were no indicators of character information past labeling the type of animal each character was (i.e., dog, cat, bird, snake, etc.); she did not use names or adjectives to describe the characters, as was taught during the MLEs. Further, her stories were similarly insufficient at pretest and posttest, in terms of the episode elements and

structure necessary to tell complete stories. Both stories included one complete episode (i.e., an initiating event, the character's attempt to respond to the initiating event, and a consequence or resolution of the initiating event), but provided limited detail surrounding such episodes. At pretest, she included one such detail: she noted a character's internal response to the initiating event ("And then the dog is still sleeping and the cat is mad"). Likewise, at posttest, she provided one similar detail indicating the character's reaction to the story's conclusion: "And then he found the ring. He was happy."

Examining the responses of another student in the typically-developing group provides an example of a competent language learner. In particular, one such student's pretest story included much less detail and complexity than her posttest story, an improvement in targeted (story components) and non-targeted (story ideas and language) areas. First, while there were some simple demonstrations of causal relationships in the pretest story, the student's posttest story explained much more clearly and frequently why the characters acted as they did; this resulted in improved scores in story components. At pretest, the student stated, "And he went away to find the cat," a statement that indicates a causal relationship, but with limited detail. Conversely, at posttest, the student made statements such as, "He was so happy that he found a ring because he thought that the ring was pretty," which provided more detailed, clear causal indicators. Further, this student's story at posttest was much more creative, resulting in improvements in story ideas and language. While, at pretest, her story did not contain much detail past those included in overt illustrations, (e.g., characters' colors, characters' basic feelings as evidenced by their illustrated facial expressions), her posttest story was far more interesting and inferential in nature. For example, she described the manner in which the

main character lost his ring as follows: “And he went and told the momma bird, “If you can save the ring?” and the momma bird said, ‘Yes.’ The momma bird was lying. She just wanted to keep the ring. So the momma bird set him up. And she changed the conversation and he was going without the ring.” This explanation goes far beyond the obvious details represented in the book’s pictures and adds to the story’s overall creativity and complexity.

To summarize, according to dynamic assessment, 88.00% of the 25 total students in the study could be categorized as typically-developing language learners, 8.00% could be deemed at-risk for SLI classification, and 4.00% were inconclusive. Conversely, the WMLS-R results indicated that 72.00% were typically-developing language learners, while 28.00% were at risk for SLI classification. Overall, according to the dynamic assessment procedure, most students in this study have typical language abilities, including all 7 identified as at-risk by their scores on the Oral Language—Total. This is an especially notable result because it differs from that of the standardized measure, which found significant differences in language ability between the at-risk and typically-developing groups. Further examples of student responses can be found in Appendix A.

CHAPTER V

DISCUSSION

Historically, educators have had significant difficulty assessing and addressing the needs of culturally and linguistically diverse (CLD) learners, especially in relation to special education classification. One important group within the CLD population is Hispanic students, who are growing in number within the United States population and who have a unique profile of needs. Educators appear to be particularly unprepared to competently assess this population's language abilities. Dynamic assessment is one method that has been suggested to be a more culturally competent procedure for assessing language ability in CLD students, including the Hispanic population. The purpose of this study was to compare a more traditional means of language assessment (i.e., the WMLS-R, a standardized measure of language) to a dynamic assessment procedure in identifying students at risk for the Speech or Language Impairment (SLI) classification. As is consistent with previous dynamic assessment studies (e.g., Pena et al., 2006; Pena, Iglesias, & Lidz, 2001), the current results indicate that solely using a traditional, standardized measure of language to assess Hispanic bilingual students for SLI may underestimate their language abilities and, thus, increase the risk for special education classification within this category. Further, dynamic assessment appears to be a valuable tool for discerning language difference from true language disorder in Hispanic children.

Language Abilities According to a Standardized Measure of Language

According to a standardized measure of language (i.e., the WMLS-R), the students in this study could be separated into two distinct groups: those with low language abilities and those with typical language abilities. Students classified with low language abilities scored below 85 on the Oral Language—Total, while students with typical language abilities scored above 85. Moreover, those with markedly lower language abilities might be at risk for SLI classification in schools, as they all met a criterion often employed by school systems to make this judgment (i.e., use of a cutoff score, as detailed in Chapter II; Oetting, Cleveland, & Cope, 2008; Spaulding, Plante, & Farinella, 2006). Also notable, is that the students at risk for SLI classification garnered significantly lower CALP (cognitive academic language proficiency) scores than the typically-developing children. This indicates that students deemed at risk for SLI classification based on the WMLS-R were also significantly less proficient with the English language than those who scored within the expected range on the WMLS-R. Specifically, Alvarado, Ruef, and Schrank (2005) state that students scoring in the range of the at-risk group are likely to have difficulty navigating the language demands of typical academic tasks, while those with scores similar to the typically-developing students would be able to do so with ease. The CALP scores were derived from the Broad English Ability—Total of the WMLS-R, which is comprised of all seven subtests. This includes the measures of expressive and receptive language also comprising the Oral Language—Total, along with measures of letter-word knowledge, written expression, and written language comprehension. Thus, while both CALP and language ability were

derived from the WMLS-R, these scores were comprised of subtests measuring different skills.

Language Abilities According to Dynamic Assessment

Following the dynamic assessment procedure, the students' narratives were scored as outlined by Miler, Gillam, and Pena (2001). Total story scores were determined by summing the story components, story ideas and language, episode elements and structure, and listener effort categories. Productivity measures included number of words, number of clauses, number of C-units, mean length of C-units, and number of clauses per C-unit. As was hypothesized, both the at-risk and typically-developing students improved significantly from pretest to posttest, both on total story scores and measures of productivity during a dynamic assessment procedure. Specifically, the at-risk group made significant improvements on all but two productivity measures (i.e., number of C-units and number of clauses per C-unit). They showed improvement in the number of words, number of clauses, and mean length of C-units. The typically-developing group showed significant gains on all but one productivity measure (i.e., number of clauses per C-unit). This group showed improvement in the number of words, number of clauses, mean length of C-units, and number of C-units. Consequently, after a dynamic assessment procedure consisting of two 30-minute mediated learning experiences, the children in this study were able to utilize language to tell more complete and complex stories.

According to the literature, when compared with typically-developing children, one would expect those with SLI to make fewer gains during dynamic assessment and to have more difficulty showing improvements during mediation, as evidenced by attaining lower modifiability ratings (Miller, Gillam, & Pena, 2001; Pena et al., 2006; Pena,

Iglesias, & Lidz, 2001; Pena, Quinn, & Iglesias, 1992). Therefore, the current study hypothesized that the children with low language abilities (i.e., the at-risk group) would perform worse and receive lower modifiability ratings than those with typical language abilities (i.e., the typically-developing group). Contrary to what was predicted, the students characterized as at risk for SLI classification demonstrated no fewer gains from pretest to posttest than the children characterized as typically-developing; both groups made significant improvements (in Total Story Scores and productivity measures) and they did so to a comparable degree. Further, children in both groups were equally able to learn and demonstrate language improvements during teaching sessions (i.e., there were no significant between-group differences in modifiability ratings). It appears that the judgments about language ability made from the standardized measure of language and those made following dynamic assessment were inconsistent. While the WMLS-R indicated clear distinctions in language ability among the students, dynamic assessment judged the students to be, as a whole, typically-developing language learners.

In order to further clarify this finding, Miller, Gillam, and Pena's (2001) dynamic assessment procedure for assessing language abilities was applied to each individual in the study. These results suggested that the overwhelming majority (i.e., 88% or 22 of 25) of students in this study were, according to dynamic assessment, typically-developing language learners, including all those classified as at risk for SLI via the standardized measure of language. Once again, this result very much differed from that of the WMLS-R, which clearly classified one discrete group of students as at risk for SLI.

Traditional Assessment Procedures Versus Dynamic Assessment

One might hypothesize that the WMLS-R and dynamic assessment judged the students' abilities differently because they were measuring different aspects of language. For example, it might be argued that the WMLS-R is a measure of CALP, while dynamic assessment is a measure of BICS. This, however, appears to be unlikely, as the ability to use and comprehend narratives is thought to be critical to make academic progress and, thus, may likely be an indicator of CALP (Gillam, McFadden, & van Kleeck, 1995; Howard, 1991; Kayser, 1995; Miller, Gillam, & Pena, 2001; Pena et al., 2006). More importantly, BICS and CALP are measures of language proficiency in a particular language rather than inherent language ability. In order to determine if a student has SLI, the evaluator must assess the underlying cognitive processes related to language ability. In an attempt to do so, dynamic assessment is focused not on static scores, but on how the child engages and progresses through the process.

Possibly a better explanation for the variation in the WMLS-R and dynamic assessment is that they seem consider the impact of culture on language very differently. Traditional school-based SLI assessment procedures have been criticized for being ill equipped for discerning language differences from language disorders in culturally and linguistically diverse students (Gillam, Pena, & Miller, 1999; Marbley, Bonner, & Berg, 2008; Miller, Gillam, & Pena, 2001; Rhodes, Ochoa, & Ortiz, 2005; Shames & Anderson, 2002). This seems to be especially concerning for Hispanic learners, as these children's differences in language experiences or styles, limited English proficiency, or lack of exposure to mainstream language learning opportunities may result in lower than expected or deserved scores on standardized measures of language ability (Artiles &

Trent, 1994; Gutierrez-Clellen & Pena, 2001; Harry, 1994; Pena & Quinn, 1997; Pena, Quinn, & Iglesias, 1992). When interpreting the results of the current study, one must consider that such factors may have played a role in the inconsistencies discovered between the standardized measure of language assessment and the dynamic assessment of language abilities. It is possible that the WMLS-R scores provided an indicator of cultural experience and/or English proficiency rather than a gauge of true language ability. This is supported by the lack of correlation between the WMLS-R Oral Language—Total scores and the dynamic assessment scores, which signifies that they were measuring different language skills. It is further substantiated by the differences in CALP scores between what the WMLS-R judged to be the at-risk and typically-developing groups. The at-risk group had significantly lower CALP scores than the typically-developing group, indicating that they were notably less proficient in English. This quality, among others related to cultural background and experiences, may have deflated this group's WMLS-R language scores and, thus, led to an inaccurate determination of their language abilities using this method. This is consistent with past studies that cautioned that static, standardized measures of language may underestimate the true language abilities of Hispanic children (Pena et al., 2006; Pena, Iglesias, & Lidz, 2001). The dynamic assessment procedure in the current study was able to, when applied to each individual, demonstrate that the students could capably learn language skills and apply teacher instruction to produce more complex and complete stories following mediation. This result suggests that students who did so were typically-developing, competent language learners; students who could not improve following mediation and who required great effort to improve during mediation would be those classified as having language learning

difficulties and, thus, at risk for SLI (Guitierrez-Clellen & Pena, 2001; Lidz & Pena, 1996; Miller, Gillam, & Pena, 2001; Pena et al., 2006; Pena, Iglesias, & Lidz, 2001; Pena, Quinn, & Iglesias, 1992).

Implications for the Use of Dynamic Assessment in SLI Classification Procedures

The implications of these results are important. First, as has been demonstrated in numerous studies, students from culturally and linguistically diverse backgrounds appear to be susceptible to the procedural inefficiencies of the special education classification process, including within the category of SLI. Standardized measures that do not control for cultural experience or bilingualism are routinely used in schools as important indicators of language ability, potentially affecting the accuracy of their results with CLD students (Oetting, Cleveland, & Cope, 2008; Rhodes, Ochoa, & Ortiz, 2005; Spaulding, Plante, & Farinella, 2006). For example, the Hispanic students in this study deemed at risk for SLI classification via one such standardized measure later demonstrated their ability to learn and utilize language during a dynamic assessment procedure. The use of inadequate evaluation tools may be one important factor impacting the disproportionate placement of CLD students in special education. It is critical that the professionals conducting SLI assessments with Hispanic students (e.g., speech-language pathologists, other Child Study Team members) be aware of the limitations of traditional procedures, be trained to delineate language differences from language disorders in these learners, and routinely incorporate these more culturally sensitive methods into SLI assessments. It appears especially critical to measure the language learning potential of CLD students, which was accomplished in the current study through just two 30-minute MLE sessions.

Notably, it does not yet seem reasonable to suggest that standardized measures be completely eliminated from the SLI assessment process; IDEIA (2004) stresses the use of technically sound evaluation instruments, standardized instruments are less time consuming, and they are aligned with one theory of SLI suggesting that children with language impairment should score lower on measures of language than do typically-developing children (Guitierrez-Clellen & Pena, 2001; Pena et al., 2006; Spaulding, Plante, & Farinella, 2006). Also, in the current study, the WMLS-R and the dynamic assessment procedure similarly judged the language abilities of 60% of the students, all of whom performed within the average range on the WMLS-R. Rather, it seems that dynamic assessment may be a useful supplement to more traditional SLI classification procedures, especially when the student of interest is from a culturally or linguistically diverse background. In this study, dynamic assessment provided the Hispanic students deemed at risk for SLI by the WMLS-R the opportunity to demonstrate their language potential in a seemingly less culturally-biased manner. For example, one student scored nearly 2 standard deviations lower than the mean on the Oral Language—Total, clearly marking him as at risk for SLI classification, and was also deemed to have limited English cognitive-academic language proficiency. However, upon participating in the dynamic assessment process, the examiner found him to be highly responsive to language instruction. He required low intensity intervention to learn and he improved substantially from pretest to posttest; his story at posttest contained more details that allowed for the listener to follow it much easier. For example, while at pretest this student began his story with no introduction or detail (“They was talking.”), at posttest he presented a much clearer picture of the story’s setting and background information related to future events:

“Once upon a time, when it was sunny, they were in the park and they found a ring.” This student was able to take limited language intervention and demonstrate noticeable improvements. Had his language abilities been judged solely on his performance on the WMLS-R, a far less meaningful portrait of his current functioning and future potential would have been produced. By utilizing dynamic assessment following the standardized assessment, the examiner was able to ascertain the student’s language learning potential and then could use this information to further refine and expand upon the results of the WMLS-R. Specifically, one might hypothesize that this student’s low scores on the standardized measure might have resulted not from a true language disorder, but from limited English proficiency. This has implications both for special education classification and future intervention.

Also, in the current study, the WMLS-R indicated that two students who scored in the low end of the typically-developing group may have truly been at risk for SLI based on their difficulty utilizing language and responding to language intervention; dynamic assessment may have picked up on language difficulties that the WMLS-R was not sensitive to. These students did not show improvement in either targeted or non-targeted areas and demonstrated only limited gains during MLEs, requiring high intensity intervention to do so. Both students told stories lacking detail, complexity, and clarity at both pretest and posttest; the stories were difficult for the listener to understand and there was minimal character and storyline development. The stories were also qualitatively very choppy and read more as disintegrated lists rather than as one cohesive story (e.g., “The horse and the cat was talking. And then the cat met the horse. And then the horse was sleeping...”). While these students demonstrated average scores on the WMLS-R,

their difficulties with language during the dynamic assessment process may indicate that further assessment is required to truly understand their educational needs. Because the required task within the dynamic assessment procedure is more typical of language usage in children's natural environments, it might be particularly important to gain a better understanding of how these students utilize language in their day-to-day settings, whether they have difficulties doing so, and where such difficulties stem from. In such cases, dynamic assessment might supplement a standardized language measure by recognizing important indicators that the WMLS-R cannot, notably, that there might be other areas of need related to language to further examine.

Implications for School Psychologists

The results of this study have clear implications for school psychologists, as they routinely participate in the classification of students with disabilities (Bramlett, Murphy, Johnson, Wallingsford, & Hall, 2002; IDEIA 2004). As an integral part of the team that determines a child's eligibility for special education, the school psychologist must be cognizant of all factors potentially impacting the child's functioning in school. As is suggested by this study and many others (e.g., Artiles & Trent, 1994; Gutierrez-Clellen & Pena, 2001; Harry, 1994), one such important factor is cultural background. It is especially imperative to examine culture when evaluating students for disabilities, such as SLI, as traditional assessment tools may not produce accurate profiles of student strengths and needs (Gillam, Pena, & Miller, 1999; Marbley, Bonner, & Berg, 2008; Miller, Gillam, & Pena, 2001; Shames & Anderson, 2002). Such inaccuracies may then lead to inappropriate special education classification (MacMillan & Reschly, 1998; Reid & Knight, 2006) and, thus, denied access to necessary educational services (Poon-

McBrayer & Garcia, 2000). In the case of SLI, school psychologists are not the primary evaluators in terms of assessing language abilities; the speech-language pathologist would likely conduct the dynamic assessment procedure. However, this should not diminish the role of school psychologists in ensuring that the evaluation is comprehensive and that all team members' assessment results are integrated in a meaningful, accurate way. Furthermore, these professionals have an obligation to see that the evaluation leads to appropriate educational services for the child. School psychologists are trained in data based decision making and are, thus, expected to be leaders in problem solving and assessment (Ysseldyke et al., 1997). This includes being able to accurately gauge student strengths and needs, provide guidance about educational programming, and evaluate the success of such programs. Moreover, they are legally and ethically compelled to work with students from diverse backgrounds in a nonbiased manner (IDEIA, 2004; Ysseldyke et al., 1997). In the case of CLD students being evaluated for SLI, it seems that dynamic assessment is a tool that can assist school psychologists with this important problem-solving process and, critically, do so fairly and informatively.

There has also been support for the use of dynamic assessment procedures to aide in the classification of children with learning disabilities (e.g., Moore-Brown, Huerta, Uranga-Hernandez, & Pena, 2006; Swanson, 1996; Swanson & Howard, 2005). The expansion of dynamic assessment to other disability categories appears particularly relevant to school psychologists, as performing cognitive assessments is a primary role of these professionals (Bramlett, Murphy, Johnson, Wallingsford, & Hall, 2002; IDEIA 2004). Cognitive assessment practices for CLD students have been historically scrutinized due to the reliance on their use for classification and the subsequent

disproportionate placement of CLD students in special education, including within the category of learning disability (NCCRESt, 2006; Oswald, Coutinho, Best, & Singh, 1999). Just as for students with suspected language impairment, it has been suggested that incorporating dynamic assessment into one's cognitive assessment battery may be particularly valuable for students from diverse cultural backgrounds; its utility, similarly, lies in determining learning potential rather than providing a one-time indicator of performance based on a measure with inherent cultural biases (Haywood & Lidz, 2007; Moore-Brown, Huerta, Uranga-Hernandez, & Pena, 2006; Saenz & Huer, 2003; Skiba, Knesting, & Bush, 2002). In particular, Moore-Brown, Huerta, Uranga-Hernandez, and Pena suggest that including dynamic assessment in learning disability evaluations can be especially helpful for achieving a more comprehensive understanding of the student and promoting a collaborative assessment process for the team members. These authors also found dynamic assessment to be especially useful for making conclusions about students whose profiles were somewhat uncommon; that is, dynamic assessment may aide the team in making classification decisions about borderline students. These results point toward the utility of dynamic assessment for a variety of reasons within the field of special education. Though the current study's results cannot be generalized to support the use of dynamic assessment with disabilities other than SLI, school psychologists must consider the potential role of dynamic assessment in a broader evaluation context.

Implications for the Use of Dynamic Assessment in Developing Interventions

Not only does dynamic assessment appear to have utility within the SLI classification process, but it also has the potential to provide more practical clinical data relative to intervention (Laing & Kamhi, 2003; Pena et al., 2006). Haywood and Lidz

(2007) assert that dynamic assessment not only provides information about current level of performance, future potential, and how well the child responds to intervention, but also about the nature of the intervention necessary to produce improvements. As a result, intervention can be expressly targeted to the mode of intervention that most readily produced language growth during mediation (Hasson & Joffe, 2007; Miller, Gillam, & Pena, 2001). Data can be gathered during MLEs, including observations of problem-solving strategies, responses to adult guidance and feedback, and ability to sustain attention to task, all of which are relevant to intervention planning (Laing & Kamhi, 2003; Pena et al., 2006). The assessment data might also be useful for developing Individualized Education Program (IEP) goals; areas of difficulty during the assessment process might indicate areas of need to be addressed by the IEP.

Miller, Gillam, and Pena (2001) suggest that developing intervention through dynamic assessment is a straightforward extension of the assessment process. However, when the professional shifts from assessment to intervention, the nature of the goals and objectives must also evolve; in intervention, the professional is focused on developing the child's skills and the child's growth as an active learner in the process. The professional utilizes the data collected during assessment to decide where to begin teaching and what will likely be the most effective manner of doing so. The professional might choose to begin teaching with an area of strong need, as evidenced in dynamic assessment scores, by using the mediation characteristics (e.g., intention to teach, meaning, transcendence, planning, and transfer) that were most effective during MLEs. One-on-one intervention sessions with the professional look similar to the mediated teaching sessions during assessment, albeit with the new goals noted previously. As a result, intervention sessions

may focus on specific areas of need for longer periods of time than during assessment (i.e., for more than two sessions). Miller, Gillam, and Pena recommend that ongoing data be collected detailing student growth, what strategies are most effective, the intensity of effort required to produce change, and the child's responsiveness to teaching. Such data can be used to demonstrate improvement and to make changes in intervention in the absence of student growth.

Though much of Millar, Gillam, and Pena's (2001) focus is on individual intervention sessions, dynamic assessment may also have implications for classroom strategies for students with SLI. The intervention goals would remain the same, while the setting within which implementation occurs would be expanded to include the classroom. In this case, individual sessions could be supported by the use of similar strategies by classroom teachers. This would require professionals (e.g., speech-language pathologists, school psychologists) to consult with teachers regarding which teaching strategies are likely to be most effective with particular students. One advantage of dynamic assessment is that teaching strategies shown to be effective in the short-term (i.e., during the assessment process) can be incorporated into classroom intervention with some confidence that they may be effective.

Using Miller, Gillam, and Pena's (2001) dynamic assessment procedure of narrative abilities may be especially valuable to a wide range of education professionals due to the pervasive nature of language. That is, because one must be competent in language to grow cognitively, academically, and socially, promoting narrative language abilities is crucial, especially for students with SLI (Gillam, McFadden, & van Kleeck, 1995; Miller, Gillam, & Pena, 2001). Students who are unable to understand or utilize

narratives may have difficulty doing so as a tool for developing cognitive and academic skills, including reading, writing, and abstract thought (Gillam, McFadden, & van Kleeck, 1995; Howard, 1991; Kayser, 1995; Miller, Gillam, & Pena, 2001). Moreover, narrative competence is important for social development, as story telling and understanding are integral parts of both interacting with others (Gillam, McFadden, & van Kleeck, 1995; Miller, Gillam, & Pena, 2001) and moderating one's own behavior (e.g., self-talk, self-regulation; Gillam, 1997).

The implications for service delivery are clear: dynamic assessment and intervention of narrative abilities is likely relevant for a variety of school professionals working with students with SLI. Speech-language pathologists will play a prominent role in utilizing dynamic assessment and intervention, as these professionals are responsible for language assessment during the classification process. Therefore, it would be the results of their dynamic assessment that would be used to generate information about language ability/potential and promising language interventions. This would most likely be used during individual sessions with speech-language pathologists, but might also inform intervention in other affected areas. The Child Study Team would have to determine which domains language difficulties are impacting and bring together the relevant professionals. All those involved would then need to collaborate in order to determine how best to carry out the intervention process, with particular guidance from the speech-language pathologist. General and special educators are those responsible for addressing the academic manifestations of students' language impairment and, thus, would deliver the necessary interventions within the classroom/academic context. Further, school psychologists and counselors would likely be responsible for promoting

the social-emotional growth of students with SLI and, thus, might deliver necessary social intervention during individual counseling sessions or social skills groups. School psychologists also often provide consultative services to teachers of students with disabilities, especially when difficulties arise. Therefore, they must be able to consult with teachers regarding the use of dynamic intervention.

Above all, expanding dynamic assessment to intervention appears to require a great deal of training and collaboration between professionals. Due to the pervasive nature of language, implementing a comprehensive dynamic intervention affects service delivery for a number of educational professionals. Speech-language pathologists could simply incorporate dynamic assessment and intervention into their preexisting individual student sessions. In regards to expansion, at the very least, teachers, school psychologists, and counselors must be instructed by speech-language pathologists regarding what type of intervention is likely to produce student growth and in what particular domains. This could be part of regular consultation that takes place between speech-language pathologists and teachers, school psychologists, and counselors; it would likely not require more of a time commitment than is already allotted. Optimally, however, these professionals would be trained in developing context-specific MLEs that effectively produce dynamic teaching sessions. While following Miller, Gillam, and Pena's (2001) and Pena et al.'s (2006) procedures is relatively straightforward, creating and implementing one's own MLEs would require much more thorough training and time. Finding this time and professionals with such expertise to conduct trainings might be draining on school resources. With limited individuals trained in dynamic assessment and

intervention, it remains to be seen how these procedures can be fully integrated into educational programming of students with SLI.

Limitations

Though the results of the current study appear practically significant, there are several limitations that must be considered. First, only twenty-five students participated in this study. Due to this small sample size, the results must be interpreted cautiously. This is especially true in light of the conclusions that may be made from the non-significant results in this study (i.e., that there were no differences in modifiability ratings and gain scores between the at-risk and typically-developing groups). In particular, it may not have been possible to detect significant differences in modifiability and gain scores among the groups had they been present within the population; limited statistical power related to small sample size accounts for such uncertainties (Cohen, 1977). The current study also did not incorporate a control condition, preventing one from definitively concluding that the students' improvements from pretest to posttest resulted from the dynamic assessment procedure rather than from inevitable growth over time, increased comfort with the testing situation, or some other factor. In particular, the students worked with the evaluator five times throughout the process, likely allowing for a steady growth in rapport and evaluator-student relationship. One might hypothesize that pretest to posttest gains were demonstrated due to the relationship and level of comfort developed between the child and the evaluator as the process progressed. It should be noted, however, that these results are consistent with similar, larger-scale, more precisely controlled studies in the literature (e.g., Pena et al., 2006; Pena, Iglesias, & Lidz, 2001). Not only did such studies show that children exposed to dynamic assessment improved

significantly more from pretest to posttest than those who did not participate in MLEs, but they also demonstrated that students with low language ability improved less than students with typical language ability.

Another methodological limitation of this study was the absence of a second individual to score the students' narratives. This made it difficult to judge the consistency and accuracy with which the scoring system was implemented. The examiner utilized Miller, Gillam, and Pena's (2001) manual in order to help maintain consistent and precise scoring procedures.

Next, the use of the WMLS-R as the traditional, standardized measure of language presents some grounds for caution. The WMLS-R is not a widely used instrument for assessing SLI, though it is composed of measures of expressive and receptive language, as is standard in SLI assessment. Further, IDEIA (2004) mandates that multiple assessment tools be utilized to determine special education eligibility. Such assessment procedures may include interviews, observations, narrative tasks, and other standardized measures (e.g., Connecticut State Department of Education, 2008; Kaderavek & Sulzby, 2000; McCabe, Bliss, Barra, & Bennett, 2008; Tyler & Tolbert, 2002; Wisconsin Department of Public Instruction, 2003). In fact, to be eligible for special education under the classification of SLI, students must not only have a language impairment demonstrated through the use of multiple assessment measures, but this impairment must impact their academic functioning. The decision was made to solely utilize the WMLS-R for a few reasons. First, though the use of other tools is required, standardized instruments play a substantial role in the classification of students with SLI in schools (Oetting, Cleveland, & Cope, 2008; Spaulding, Plante, & Farinella, 2006).

Similar procedures as that used in this study have also been utilized to screen culturally and linguistically diverse students for SLI (e.g., Tyler & Tolbert, 2002). Therefore, while this procedure did not strictly mirror mandated SLI assessment procedures, it included one aspect emphasized in schools and corresponded with SLI screening procedures. As such, labeling the students who scored low on the WMLS-R as at risk for SLI classification rather than definitively classifiable seems particularly appropriate.

Finally, this study did not measure the students' proficiency with the Spanish language. Doing so might have yielded further insight into the students' language profiles and helped further refine and classify their language abilities. This is especially noteworthy due to the differences in English cognitive-academic language proficiency between the at-risk and typically-developing groups. This and their status as bilingual, Hispanic-Americans indicates that they might have varying proficiency with the Spanish language. Evaluating their Spanish language proficiency could have provided further support for the hypothesis that limited English language proficiency and limited exposure to the mainstream culture negatively impacted WMLS-R performance for the at-risk students. Lack of proficiency in either language, or semilingualism, but ability to make gains during dynamic assessment might further support the claim that language difficulties were not due to language impairment but to lack of exposure to the mainstream language. Other profiles might support the opposite hypothesis: that students have SLI. One such profile might be a student with adequate English language proficiency, very limited Spanish language proficiency, and inability to demonstrate improvements during dynamic assessment. Overall, the inclusion of Spanish language

assessments may have allowed for more complete judgments to be made about language proficiency, cultural background, and language ability.

Future Research and Conclusions

Future research should expand upon these results by further comparing more complete traditional SLI assessment procedures with dynamic assessment of language ability. More complete procedures would include more comprehensive measures of language, multiple methods of assessment, and evidence of academic performance. They may also include bilingual assessments for students who speak languages other than English. The goal should be to further understand the role of dynamic assessment in SLI assessment, whether that is to discern language differences from language disorders in CLD students, to be included as part of a standard SLI assessment battery, to inform intervention, and/or to serve other purposes. More research is also needed regarding the long-term effectiveness of teaching strategies utilized during the dynamic assessment process. In particular, are the teaching strategies deemed effective during MLEs truly effective in producing long-term development in children with SLI?

This study found that a standardized measure of language (i.e., WMLS-R) and a dynamic assessment procedure differed with respect to their judgments of language ability in Hispanic bilingual first graders. Students with low scores on the WMLS-R (i.e., 1 standard deviation below the mean) were deemed at risk for SLI classification, while those scoring at or above average were classified as typically-developing. Both the students at risk for SLI classification and the typically-developing students demonstrated significant improvements from pretest to posttest during a dynamic assessment procedure. Moreover, both groups displayed similar growth from pretest to posttest and

were rated similarly on modifiability. When the dynamic assessment procedure was applied at the individual student level, the results suggested that the majority of students were, according to dynamic assessment, typically-developing language learners, including all those classified as at risk by the WMLS-R.

These results serve as preliminary support for the use of dynamic assessment as a supplement to traditional SLI assessment procedures with Hispanic children. The results are consistent with those similar studies in the literature that indicate that evaluators may fail to attain a comprehensive understanding of student abilities and needs if they strictly rely upon standardized assessment tools as evidence of language ability. The current study indicates that Hispanic students may be particularly vulnerable to this assessment misstep and, more importantly, that dynamic assessment may be helpful in ameliorating the cultural unfairness of traditional procedures. Above all, it is imperative to promote growth among educators in alignment with the continuously evolving United States population. Developing culturally competent professionals who have the tools to evaluate the abilities and needs of diverse children, will be a central factor in the drive to foster the educational success of all children.

REFERENCES

- Alvarado, C. G., Ruef, M. L., & Schrank, F. A. (2005). Comprehensive Manual. *Woodcock-Munoz Language Survey—Revised*. Itasca, IL: Riverside Publishing.
- Anderson-Yockel, J. & Haynes, W. O. (1994). Joint book-reading strategies in working-class african american and white mother-toddler dyads. *Journal of Speech and Hearing Research*, 37, 583- 593.
- Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2002). English language learner representation in special education in california urban school districts. In D. J. Losen & G. Orfield (Eds.), *Racial inequity in special education* (pp. 117-135). Cambridge, MA: Harvard Education Publishing Group.
- Artiles, A. J., & Trent, S. C. (1994). Overrepresentation of minority students in special education: A continuing debate. *Journal of Special Education*, 27, 410-437. doi:10.1177/002246699402700404
- Bashir, A. S., Conte, B. M., & Heerde, S. M. (1998). Language and school success: Collaborative challenges and choices. In D. Merritt & B. Culatta (Eds.), *Language intervention in the classroom* (1-36). San Diego, CA: Singular Publishing Group.
- Bedore, L. M. & Pena, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11, 1-29.
- Bloom, L. (1970). *Language development: Form and function of emerging grammars*. Cambridge, MA: MIT Press.
- Bloom, L. & Lahey, M. (1978). *Language development and language disorders*. New York, NY: Wiley.
- Bramlett, R.K., Murphy, J.J., Johnson, J., Wallingsford, L. & Hall, J.D. (2002). Contemporary practices in school psychology: A national survey of roles and referral problems. *Psychology in the Schools*, 39, 327-335. doi:10.1002/pits.10022
- Budoff, M. (1987). Measures for assessing learning potential. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 173-195). New York, NY: The Guilford Press.
- Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, 3, 1-61.
- Chomsky, N. (1975). *The logical structure of linguistic theory*. Chicago, IL: University of Chicago Press.

- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York, NY: Academic Press.
- Connecticut State Department of Education. (2008). *Guidelines for speech and language programs: Determining eligibility for special education speech and language services under IDEA*. Middletown, CT: Connecticut State Department of Education.
- Cummins, J. (1984). *Bilingualism and special Education: Issues in assessment and pedagogy*. San Diego, CA: College-Hill.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 222-251.
- Cummins, J. (1981). *The role of primary language development in promoting educational success for language minority students: A theoretical framework*. Los Angeles, CA: National Dissemination Center.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59 (no. 5), Serial No. 242.
- Feuerstein, R. (1979). *Dynamic assessment of retarded performers*. Baltimore, MD: University Park Press.
- Feuerstein, R. (1980). *Instrumental enrichment*. Baltimore, MD: University Park Press.
- Feuerstein, R., Rand, Y., Jensen, M. R., Kaniel, S., & Tzuriel, D. (1987). Prerequisites for assessment of learning potential: The LPAD model. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp.35-51). New York, NY: The Guilford Press.
- Fey, M. (1986). *Language intervention with young children*. Austin, TX: Pro-ed.
- Fletcher, T. V. & Navarrete, L. A. (2003). Learning disabilities or difference: A critical look at issues associated with the misidentification and placement of hispanic students in special education programs. *Rural Special Education Quarterly*, 22, 37-46.
- Gillam, R. B. (1997). Putting memory to work in language intervention: Implications for practioners. *Topics in Language Disorders*, 18, 72-79. doi:10.1097/00011363-199711000-00008
- Gillam, R. B., McFadden, T. U., & van Kleeck, A. (1995). Improving the narrative abilities of children with language disorders: Whole language and language skills

- approaches. In M. Fey, J. Windsor, & J. Reichle (Eds.), *Communication intervention for school-age children* (pp. 145-182). Baltimore, MD: Brookes.
- Gillam, R. B., Pena, E. D., & Miller, L. (1999). Dynamic assessment of narrative and expository discourse. *Topics in Language Disorders*, 20, 33-47. doi:10.1097/00011363-199911000-00005
- Gleitman, L. R. & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary development. *Current Directions in Psychological Science*, 1, 31-35. doi:10.1111/1467-8721.ep10767853
- Gottlieb, J., Alter, M., Gottlieb, B. W., & Wishner, J. (1994). Special education in urban america: It's not justifiable for many. *Journal of Special Education*, 27, 453-465. doi:10.1177/002246699402700406
- Guiberson, M. (2009). Hispanic representation in special education: Patterns and implications. *Preventing School Failure*, 53, 167-176. doi:10.3200/PSFL.53.3.167-176
- Gutierrez-Clellen, V. & Iglesias, A. (1987, November). *Expressive vocabulary of kindergarten and first grade hispanic students*. Paper presented at the American Speech-Language-Hearing Association national convention, New Orleans.
- Gutierrez-Clellen, V. F. & Pena, E. (2001). Dynamic assessment of diverse children: A tutorial. *Language, Speech, and Hearing Services in Schools*, 32, 212-224. doi:10.1044/0161-1461(2001/019)
- Harry, B. (1994). *The disproportionate representation of minority students in special education: Theories and recommendations* (Project FORUM). Alexandria, VA: National Association of State Directors of Special Education.
- Hasson, N. & Joffe, V. (2007). The case for dynamic assessment in speech and language therapy. *Child Language, Teaching, and Therapy*, 23, 9-25. doi: 10.1177/0265659007072142
- Haywood, H. C. & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. New York, NY: Cambridge University Press.
- Hearne, D. (2000). *Teaching second language learners with learning disabilities: Strategies for effective practice*. Oceanside, CA: Academic Communication Associates.
- Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society*, 11, 49-76. doi:10.1017/S0047404500009039

- Heath, S. B. (1986). Sociocultural contexts of language development. In California State Department of Education (Ed.), *Beyond language: Social and cultural factors in schooling language minority children* (pp. 143-186). Los Angeles, CA: Evaluation, Dissemination and Assessment Center, California State University.
- Hosp, J. L., & Reschly, D. J. (2003). Referral rates for intervention or assessment: A meta-analysis of racial differences. *The Journal of Special Education*, 37, 67-80. doi:10.1177/00224669030370020201
- Howard, G. (1991). Culture tales: A narrative approach to therapy in cross-cultural psychology and psychotherapy. *American Psychologist*, 40, 187-197. doi:10.1037/0003-066X.46.3.187
- Individuals with Disabilities Education Improvement Act of 2004, P.L. 108-446. Retrieved from <http://idea.ed.gov/download/statute.html>
- Jusczyk, P. W. (1995). Language acquisition: Speech sounds and phonological development. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition* (263-301). Orlando, FL: Academic Press.
- Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A combination of vocal, dynamic, and summary features discriminates between three pragmatic categories of infant-directed speech. *Child Development*, 67, 205-217. doi:10.1111/j.1467-8624.1996.tb01729.x
- Kayser, H. (1995). Assessment of speech and language impairments in bilingual children. In H. Kayser (Ed.), *Bilingual speech-language pathology: A Hispanic focus* (pp.1-14). San Diego, CA: Singular.
- Kummerer, S. E. (2010). Language intervention for hispanic children with language-learning disabilities: Evidence-based practice. *Intervention in School and Clinic*, 45, 192-200.
- Kummerer, S. E., Lopez-Reyna, N. A., & Hughes, M. T. (2007). Mexican immigrant mothers' perceptions of their children's communication disabilities, emergent literacy development, and speech-language therapy program. *American Journal of Speech-Language Pathology*, 16, 271-282.
- Laing, S. P. & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, and Hearing Services in Schools*, 34, 44-55. doi:10.1044/0161-1461(2003/005)
- Langdon, H. W. & Cheng, L. L. (1992). *Hispanic children and adults with communication disorders: Assessment and intervention*. Gaithersburg, MD: Aspen.

- Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. New York, NY: The Guilford Press.
- Lidz, C. S. & Pena, E. D. (1996). Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to latino american preschool children. *Language, Speech, and Hearing Services in Schools*, 27, 367-372.
- MacMillan, D. L., & Reschly, D. J. (1998). Overrepresentation of minority students: The case for greater specificity or reconsideration of the variables examined. *Journal of Special Education*, 32, 15-24. doi:10.1177/002246699803200103
- Marbley, A. F., Bonner, F., & Berg, R. (2008). Measurement and assessment: Conversations with professional people of color in the field of education. *Multicultural Education*, 16, 12-20.
- Mayer, M. & Mayer, M. (1975). *One frog too many*. New York, NY: Dial Books for Young Readers.
- McLean, J. & Snyder-McLean, L. (1978). *A transactional approach to early language training*. Columbus, OH: Merrill.
- Merrell, A. W. & Plante, E. (1997). Norm-referenced test interpretation in the diagnostic process. *Language, Speech, and Hearing Services in Schools*, 28, 50-58.
- Miller, L. (1999a). *Bird and his ring*. Austin, TX: Neon Rose Productions.
- Miller, L. (1999b). *Two friends*. Austin, TX: Neon Rose Productions.
- Miller, L., Gillam, R. B., & Pena, E. D. (2001). *Dynamic assessment and intervention: Improving children's narrative abilities*. Austin, TX: Pro-ed.
- Moore-Brown, B., Huerta, M., Uranga-Hernandez, Y., & Pena, E. D. (2006). Using dynamic assessment to evaluate children with suspected learning disabilities. *Intervention in School and Clinic*, 41, 209-217. doi:10.1177/10534512060410040301
- National Center on Culturally Responsive Education Systems (NCCRESt). (2006). *Disproportionality by race and disability 2005-2006* [Data file]. Retrieved from http://nccrest.eddata.net/maps/index.php?id=47&col=RACE_RRPUB&fl=2005-2006&group=Hispanic
- New York University, School of Education. (1994). *An analysis of referrals, placement, and progress of children with disabilities who attend new york city public schools: Final report*. Bronx, NY: M. Gottlieb & J. Alter.

- Northwest Regional Educational Laboratory (2003). *Overview of second language acquisition and strategies*. Retrieved from <http://www.nwrel.org/request/2003may/overview.html>
- Oetting, J. B., Cleveland, L. H., & Cope III, R. F. (2008). Empirically derived combinations of tools and clinical cutoffs: An illustrative case with a sample of culturally/linguistically diverse children. *Language, Speech, and Hearing Services in Schools*, 39, 44-53.
- Oswald, D. P., Coutinho, M. J., Best, A. M., & Singh, N. N. (1999). Ethnic representation in special education: The influence of school-related economic and demographic variables. *Journal of Special Education*, 32, 194-206. doi:10.1177/002246699903200401
- Owens, R. E. (2005). *Language development: An introduction* (Vol. 6). Geneseo, NY: Pearson Education, Inc.
- Patton, J. M. (1998). The disproportionate representation of african americans in special education: Looking behind the curtain for understanding and solutions. *Journal of Special Education*, 32, 25-31. doi:10.1177/002246699803200104
- Pena, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., & Sabel, T. (2006). Dynamic assessment of school-age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech, Language, and Hearing Research*, 49, 1037-1057. doi:10.1044/1092-4388(2006/074)
- Pena, E., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology*, 10, 138-154. doi:10.1044/1058-0360(2001/014)
- Pena, E. D. & Quinn, R. (1997). Task familiarity: Effects on the test performance of puerto rican and african american children. *Language, Speech, and Hearing Services in Schools*, 28, 323-332.
- Pena, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. *Journal of Special Education*, 26, 269-280. doi:10.1177/002246699202600304
- Perez, B., Skiba, R. J., & Chung, C. (2008). Latino students and disproportionality in special education. *Education Policy Brief*, 6, 1-8.
- Pew Hispanic Center. (2008). *One-in-five and growing fast: A profile of hispanic public school students*. Washington, DC: R. Fry & F. Gonzales.

- Poon-McBrayer, K. F. & Garcia, S. B. (2000). Profiles of asian american students with LD at initial, referral, assessment, and placement in special education. *Journal of Learning Disabilities*, 33, 61-71. doi:10.1177/002221940003300109
- Reid, D. K., & Knight, M. G. (2006). Disability justifies exclusion of minority students: A critical history grounded in disability studies. *Educational Researcher*, 35, 18-23. doi:10.3102/0013189X035006018
- Rhodes, R. L., Ochoa, S. H., & Ortiz, S. O. (2005). *Assessing culturally and linguistically diverse students: A practical guide*. New York, NY: The Guilford Press.
- Roseberry-McKibbin, C. (2002). *Multicultural students with special language needs* (2nd ed.). Oceanside, CA: Academic Communication Associates.
- Saenz, T. I. & Huer, M. B. (2003). Testing strategies involving least biased language assessment of bilingual children. *Communication Disorders Quarterly*, 24, 184-193. doi:10.1177/15257401030240040401
- Salend, S. J., Duhaney, M. G., & Montgomery, W. (2002). A comprehensive approach to identifying and addressing issues of disproportionate representation. *Remedial and Special Education*, 23, 289-299. doi:10.1177/07419325020230050401
- Salvia, J. & Ysseldyke, J. E. (1991). *Assessment* (5th ed.). New York, NY: Houghton Mifflin.
- Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, 23, 3-12. doi:10.1037//0012-1649.33.1.3
- Schon, J., Shaftel, J., & Markham, P. (2008). Contemporary issues in the assessment of culturally and linguistically diverse learners. *Journal of Applied School Psychology*, 24, 163-189. doi:10.1080/15377900802089395
- Shames, G. H. & Anderson, N. B. (2002). *Human communication disorders: An introduction* (6th ed.). Boston, MA: Allyn and Bacon.
- Skiba, R. J., Knesting, K., & Bush, L. D. (2002). Culturally competent assessment: More than nonbiased tests. *Journal of Child and Family Studies*, 11, 61-78. doi:10.1023/A:1014767511894
- Skiba, R. J., Simmons, A. B., Ritter, S., Gibb, A. C., Rausch, M. K., Cuadrado, J., & Chung, C. (2008). Achieving equity in special education: History, status, and current challenges. *Exceptional Children*, 74, 264-288.
- Skinner, B. F. (1957). *Verbal behavior*. New York, NY: Appleton-Century-Crofts.

- Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Eligibility criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech, and Hearing Services in Schools*, 37, 61-72.
- Swanson, H. L. (1996). Classification and dynamic assessment of children with learning disabilities. *Focus on Exceptional Children*, 28, 1-20.
- Swanson, H. L. & Howard, C. B. (2005). Children with reading disabilities: Does dynamic assessment help in the classification? *Learning Disability Quarterly*, 28, 17-34. doi:10.2307/4126971
- Trawick-Smith, J. (2003). *Early childhood development: A multicultural perspective* (3rd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.
- Tyler, A. A. & Tolbert, L. C. (2002). Speech-language assessment in the clinical setting. *American Journal of Speech-Language Pathology*, 11, 215-220. doi:10.1044/1058-0360(2002/022)
- U. S. Department of Education, Office of Special Education and Rehabilitative Services. (2006). *28th annual report to congress on the implementation of the individuals with disabilities education act* (Vol. 2). Washington DC: Westat.
- Vygotsky, L. S. (1978). Interaction between learning and development. In L. S. Vygotsky, *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. and Trans.) (pp. 79-91). Cambridge, MA: Harvard University Press. (Original work published 1935).
- Vygotsky, L. S. (1986). Thinking and speech. In L. S. Vygotsky, *Collected works: Problems of general psychology* (Vol. 1) (N. Minick, Trans.). New York, NY: Plenum. Translation of V. V. Davydov (Ed.). (1982). *L. S. Vygotskii: Sobranie sochinenii* [L. S. Vygotsky: Collected Works] (Vol. 2). Moscow, RU: Pedagogika. (Original work published 1934).
- Washington, J. A. & Craig, H. K. (2004). A language screening protocol for use with young african american children in urban settings. *American Journal of Speech-Language Pathology*, 13, 329-340. doi:10.1044/1058-0360(2004/033)
- Wisconsin Department of Public Instruction. (2003). *Speech and language impairments assessment and decision making: Technical assistance guide*. Madison, WI: Wisconsin DPI.
- Woodcock, R. W., Munoz-Sandoval, A. F., Ruef, M. L., & Alvarado, C. G. (2005). *Woodcock-Munoz Language Survey—Revised, English*. Itasca, IL: Riverside Publishing.

Ysseldyke, J., Dawson, P., Lehr, C., Reschly, D., Reynolds, M., & Telzrow, C. (1997). *School psychology: A blueprint for training and practice II*. Bethesda, MD: National Association of School Psychologists.

APPENDIX A

Examples of Typically-Developing Language Learners and Those At Risk for Language Learning Difficulties During a Dynamic Assessment Procedure

A. Typically-Developing: Adequate modifiability rating and very high scores at pretest (Subject ID #1200; Classified typically developing by WMLS-R)

Pretest: A

“Once upon a time the dog asked the cat if he can have a play date with the cat. The cat said, ‘Maybe tomorrow.’ The next day, the cat went to the dog’s house. The dog was sleeping so the cat got mad. The cat was thinking if he can leave the dog because now he didn’t want to have the play date. While the dog was sleeping the cat was moving on, the dog was sleeping all day and night. When the dog woke up, the dog remembered that today was supposed to be when the dog was supposed to have a play date with the cat and the dog said, ‘I’m going to find the cat to have a play date with him.’ Then the dog marched along the road. The dog asked a dragon to see if he has seen a little cat because he was going to have a play date with her. And the big dragon said, ‘No.’ Then he came to a armadillo and asked the armadillo if he has seen a cat so he can have a play date with him and the armadillo said, ‘No.’ Then a lion was swimming in the water. The dog ran out of the water and went to see if there was any clue and then the dog found the tail of the cat. Then the dog apologized to the cat and the cat said, ‘It’s ok’ and the cat said, ‘Let’s have a play date by watching the stars.’”

Scores:

Variable	Score
Total Story Score	45.00
Number of Words	253.00

Number of Clauses	42.00
Mean Length of C-unit	12.65
Number C-units	20.00
Number of Clauses per C-unit	2.10

Posttest: A

“One summer day in Texas a bird flew down to get food. On his way he saw a cactus with a shiny ring on it so he took it. The bird was walking down the road when he passed by a snake. Then the bird flew up to her nest and gave the ring to her little baby chick. Then she went down to get more food and then she came back up, gave the food to her little yellow chick, and then she was talking to everyone about the ring her little yellow chick has. While the sun was going down and the clouds were getting gray, it looked like it was going to rain. Then he told a lizard about the baby yellow chick with the ring while the sun was going down lower and the clouds were getting pink. The lizard went up to see it. The baby chick was crying because he didn’t have it anymore. Then the bird asked everyone if he had found a ring and everyone said, ‘No.’ the mommy bird was sad. She was talked to her friend, Lizzy, about the ring and he had said, ‘No.’ The sun was going down and the sky was getting a little dark. And Lizzy said, ‘I think I found it.’ And they all went down to find it and they found the ring. The mommy bird brought it to the baby chick.”

Scores:

Variable	Score
Modifiability	10.00

Total Story Score	48.00
Number of Words	238.00
Number of Clauses	40.00
Mean Length of C-unit	10.35
Number C-units	23.00
Number of Clauses per C-unit	1.74

B. Typically-Developing: Adequate modifiability rating and significant improvements made from pretest to posttest (Subject ID #1115; Classified typically developing by WMLS-R)

Pretest: B

“One day Kate and Sam was talking. Then Sam was talking to Kate. Then Kate said a mean thing to Sam. Then Kate said, ‘Go away Sam.’ Then Kate was talking. Then Kate was thinking about Sam. Then Kate was running on the hill. Then he saw his master, Max, and Kate said, ‘I scared Max away.’ Kate was talking about Sam with his friend Pat. Then Kate was swimming in the water. Then Kate running on the grass. Then they made friends again by saying nice things to each other.”

Scores:

Variable	Score
Total Story Score	25.00
Number of Words	91.00
Number of Clauses	16.00
Mean Length of C-unit	7.00
Number C-units	13.00

Number of Clauses per C-unit	1.23
------------------------------	------

Posttest: B

“One day Max saw a ring on a cactus. Then he walked and walked and walked. Then he stopped at Kate’s nest. Then he said, ‘Here Kate. Have the ring in your nest and keep it safe because if you don’t keep it safe it will get lost.’ Then Max was walking along to his house. Then he was thinking about the ring. Then he looked up at his friend Kate. ‘Where is my ring?’ he said. ‘I don’t know,’ she said. ‘I forgot it was lost.’ ‘My ring, my ring!’ he said. ‘Where is my ring?’ He asked the lizard, ‘Have you seen my ring?’ ‘No,’ the lizard said. Then he was sad. Then he thinked and thinked and thinked and thinked. Then he went out in the forest. Then he saw a ring behind his friend’s tree. It was there. So he found the ring. He was so much happy. Now he was superdeduperdy happy. It was awesome.”

Scores:

Variable	Score
Modifiability	9.00
Total Story Score	39.00
Number of Words	160.00
Number of Clauses	31.00
Mean Length of C-unit	6.40
Number C-units	25.00
Number of Clauses per C-unit	1.24

C. Typically-Developing: Adequate modifiability rating and significant improvements made from pretest to posttest (Subject ID #1116; Classified at risk by WMLS-R)

Pretest: C

“The dog is saying hi to the cat. And the dog was sleeping. Then the cat was leaving away. And then the dog was sleeping again and the cat got sad. And the dog was sleeping again and the cat got away and away and away. And the dog was still sleeping. Then he wake up. Then he didn’t saw the cat. Then the dog was running and finding the cat. And the dog found a dragon. And then the dog saw a animal. And the dog was swimming and swimming, but the cat was not there. But the dog saw the cat’s tail. And then the dog found the cat. Then they be friends.

Scores:

Variable	Score
Total Story Score	18.00
Number of Words	114.00
Number of Clauses	19.00
Mean Length of C-unit	6.33
Number C-units	18.00
Number of Clauses per C-unit	1.06

Posttest: C

“One morning, Tom saw a ring and Tom put the ring in his mouth. Then he take it to his nest. And Tom left and the ring was in the nest shining because the sun was shining at the ring. And tom saw a iguana and Tom screamed that maybe the ring is ok. And then Tom went to the nest. Then the bird said, ‘I don’t know where the ring. I was here seeing the

ring then somebody took it.’ And then the bird screamed. And then Tom was looking around and the iguana was up the tree seeing the bird scream. And tom said to a porcupine, ‘You saw a ring around here?’ Then the porcupine said, ‘No’ and then the iguana was in the ground and moving his head. And then Tom said to the alien, ‘You saw a ring around here?’ And then the iguana said the same that Tom said too. The iguana copied Tom. And then Tom was looking around the forest and the iguana too. Then tom and the iguana found the ring and the bird didn’t saw it because the ring was under the tree. And Tom put the ring in his mouth and the bird said, ‘How did you find it?’ ‘I found it under the tree.’ And the iguana was sleeping on the tree.”

Scores:

Variable	Score
Modifiability	9.00
Total Story Score	30.00
Number of Words	224.00
Number of Clauses	37.00
Mean Length of C-unit	8.00
Number C-units	28.00
Number of Clauses per C-unit	1.32

D. At risk: Low modifiability rating and no significant improvements from pretest to posttest (Subject ID #1151; Classified typically developing by WMLS-R)

Pretest: D

“The horse and the cat was talking and then the cat met the horse and then the horse was sleeping. And then the cat was talking and the horse was still sleeping and the cat went

away. And then the horse was still sleeping. Then the horse was thinking about the cat. The bird came and was following the horse. And a horse met a bear and then the horse met an armadillo. Next the horse met fish and the horse was following the cat. Then the horse and the cat were friends.”

Scores:

Variable	Score
Total Story Score	15.00
Number of Words	93.00
Number of Clauses	15.00
Mean Length of C-unit	6.64
Number C-units	14.00
Number of Clauses per C-unit	1.07

Posttest: D

“The bird saw a ring on the cactus. Then he took the ring. Then he showed the bird. Then the bird took it. And then the bird was thinking about the ring. The bird said, ‘Where is the ring at?’” The lizard was trying to get the ring too. Then he was talking about the ring. Then he looked at the head and was talking about the ring. Then he thought the ring was there. Then they were spying on the bird. Then he got the ring back.”

Scores:

Variable	Score
Modifiability	5.00
Total Story Score	17.00
Number of Words	88.00

Number of Clauses	13.00
Mean Length of C-unit	7.33
Number C-units	12.00
Number of Clauses per C-unit	1.08
