USING THE DEVELOPMENTAL INDICATORS FOR THE ASSESSMENT OF LEARNING – THIRD EDITION AS A SCREENER FOR YOUNG CHILDREN: A COMPARISON OF THE PSYCHOMETRIC PROPERTIES BETWEEN THE ENGLISH AND SPANISH-SPEAKING STANDARDIZATION SAMPLES

A Dissertation

by

ROMAN GARCIA DE ALBA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2006

Major Subject: School Psychology

USING THE DEVELOPMENTAL INDICATORS FOR THE ASSESSMENT OF LEARNING – THIRD EDITION AS A SCREENER FOR YOUNG CHILDREN: A COMPARISON OF THE PSYCHOMETRIC PROPERTIES BETWEEN THE ENGLISH AND SPANISH-SPEAKING STANDARDIZATION SAMPLES

A Dissertation

by

ROMAN GARCIA DE ALBA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Salvador Hector Ochoa
	Michael J. Ash
Committee Members,	Victor Willson
	Victor Arizpe
Head of Department,	Michael Benz

May 2006

Major Subject: School Psychology

ABSTRACT

Using the Developmental Indicators for the Assessment of Learning – Third Edition as a Screener for Young Children: A Comparison of the Psychometric Properties between the English and Spanish-Speaking Standardization Samples. (May 2006)

Roman Garcia de Alba, B.A., Texas A&M University

Co-Chairs of Advisory Committee: Dr. Salvador Hector Ochoa
Dr. Michael J. Ash

Demographic data show that public schools are faced with meeting the academic demands of a population that is becoming more ethnically and linguistically diverse. Preventative steps can give schools the opportunity to address the needs of its students before systemic inefficiencies can negatively impact student academic outcomes. For this reason, it is important that school psychologists remain vigilant regarding the most efficient and cost effective means to identify problems early. Since Spanish is the most prevalent language of children in the schools other than English, there is a need for school psychologists to find screening instruments that are specifically designed to convey an accurate representation of the abilities of this population. One screening instrument that has been posited as effective in assessing both English and Spanish-speakers is the Developmental Indicators for the Assessment of Learning – Third Edition (DIAL-3). The purpose of this study is to expand the work of the DIAL-3 authors to include more detailed information regarding its reliability and validity for the Spanish-

speaking sample. This study was conducted using the data from the standardization samples of both the English and the Spanish versions of the DIAL-3.

Given the nature of the instrument, the obtained reliability estimates, computed using Cronbach's α, fell within the expected range. Reliability estimate comparisons between English and Spanish-speaking samples were not statistically significant with the exception of the reliability comparisons in two domains of the DIAL-3 in the 3 years 0 months to 3 years 5 months age range. Results from additional statistical analyses conducted for this study support the discriminant validity of the test. However, a moderate linear relationship was found between the Concepts and Language Domains ($r = .61$, $p < .01$). In addition, a series of confirmatory factor analyses were conducted in order to determine the invariance of the variance-covariance matrices between the English and Spanish standardization samples. The four fit indices examined (GFI, CFI, NFI, and RMSEA) for the constrained model were within the acceptable limits. These results indicate that the three-factor model originally proposed by the test authors is adequate for both the English and Spanish versions of the DIAL-3.

DEDICATION

This dissertation is dedicated to my parents
whose unconditional love, generosity and
support have been and will continue to be the
impetus behind my life's work.

ACKNOWLEDGMENTS

First and foremost, I would like to thank Dr. Salvador Hector Ochoa for being my mentor throughout graduate school. Your guidance and support in the past five years have been invaluable. Thank you for always believing in me and for molding me into the professional that I am today. I would also like to thank Dr. Michael J. Ash for helping me think "outside the box" by challenging my points of view and for teaching me about the many realities of becoming a school psychologist, including the need to relax and enjoy life. I would also like to thank Dr. Victor Willson for always being generous with his time and for being willing to guide me through the challenging world of statistics. I have learned many things from you. I would also like to thank Dr. Victor Arizpe, who I had the privilege of getting acquainted with during my undergraduate years. You were one of the first Hispanic professors I met while attending Texas A&M and showed me the importance of working with students and becoming a leader for the Hispanic community. Many thanks also go to Dr. Cynthia A. Riccio who taught me the importance of doing quality work and who also taught me many of the skills that are essential to becoming a successful school psychologist. I would like to give a special thanks to American Guidance Services, Inc., who by giving me permission to use DIAL-3 standardization data, provided me with the opportunity to conduct this study.

In addition to the people that were key in my formation as a professional, I would also like to thank Kathy May, Carol Wagner, Angela Albrecht, and Teresa Roberts who have always given me a helping hand in dealing with difficult people and situations, finances, department politics, etc. (i.e. areas that are not necessarily covered in the

graduate handbook). You all made my life at A&M more tolerable especially when the pressures of graduate school seem to get the best of me. I would also like thank Noe Ramos (a.k.a. O.P.). You have been my right-hand man in every life endeavor I have undertaken in these past few years. Thanks for all your help, advice, support and motivation. I would not be where I am now without you. Last, but certainly not least, I want to thank my family, but especially my mom and dad. You instilled in me the true value of service, dedication, and always giving your best in everything you do.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Public schools in the U.S. are continually challenged to meet the demands of the children they serve in a manner that is compliant with governing laws. Demographic data shows that public schools are faced with meeting the academic demands of a population that is becoming more ethnically and linguistically diverse. Of those children who attended schools in 1999, approximately 61% were Caucasian, 17% were Black, 16% were Hispanic, and 5% were Asian/Pacific Islander (U.S. Bureau of the Census, 1999). Linguistically, the school population has also become diverse with approximately 4.6 million children presently being classified as Limited English Proficient (Kindler, 2002).

There are factors that are associated with academically successful children. Winquist-Nord and Griffin (1999) mention that the success children experience in school is determined by those experiences that take place before the child ever enters a classroom. Hart and Risley (1992) mention that a child's optimal cognitive outcomes are associated with home environments that offer varied opportunities to learn through interaction with caring adults and age appropriate materials.

Unfortunately, some children, especially those from culturally and linguistically diverse backgrounds, do not have the privilege of relying on the aforementioned learning opportunities and/or resources. Therefore, it is also important to become familiar with the

---

This dissertation follows the style and format of the *Journal of School Psychology.*

risk factors that hinder the academic success of children. Some of these variables include poverty, the educational attainment of parents, living in a single parent household, and the number of siblings. In order to protect the welfare of children with academic needs and who endure one or more of the aforementioned factors, several laws were enacted to provide them with opportunities for success in school. Children with disabilities may be protected by three federal statutes: Section 504 of the Rehabilitation Act of 1973 (Section 504), the Americans with Disabilities Act of 1990 (ADA), and the Individuals with Disabilities Education Act, Part C (IDEA) (1997). One critical component in the school accountability process is its Child Find process. According to the Individuals with Disabilities Education Act (IDEA) (1997), schools are responsible for identifying, locating, and evaluating all preschool children who may evidence a potential disability and subsequently, are in need of early intervention or special education services (P.L. 105-17). According to Thurlow (1992), prior to 1975 a large number of children were entering school with difficulties that impacted their ability to learn as well as children that had academic problems that could have been addressed when they were younger.

Accountability for these children can easily become a problem if preventative steps are not implemented to identify problems early to be in compliance with the aforementioned federal statute. Preventative steps can give schools the opportunity to address the needs of its students before systemic inefficiencies can negatively impact student academic outcomes. Since one of the goals of IDEA is to determine the academic needs of children at an early stage, identification of these students should take place

before or at the beginning of their academic life in the public school system approximately during ages 3 to 5 or 6 years.

Approximately 1 million children between the ages of 3 to 6 years are currently being served under IDEA (U.S. Department of Education, 2003). To ease the onus placed on the schools, it is important that school psychologists and other diagnostic personnel remain vigilant regarding the most efficient and cost effective means to identify problems early. Unfortunately, due to the large numbers of children with academic needs in addition to a limited school budget, schools cannot afford to evaluate every child comprehensively when a child's competencies are in question. Therefore, an alternative way in which school psychologists can accomplish this task is through the use of screening instruments. However, assessing young children can be difficult given that they have different rates of maturation, different experiential, cultural, and linguistic backgrounds (Romero, 1999). Since these factors impact the performance of the child during an evaluation, it is important that the instruments used meet the highest standards so that their influence is minimized.

The Code of Fair Testing Practices in Education has provided specific criteria for psychologists to ensure that these standards are met (Joint Committee on Testing Practices, 2003). Knowledge of these criteria assists school psychologists in obtaining valid information regarding the preschool child. In addition to the aforementioned criteria, IDEA clearly stipulates several procedures that must be addressed when evaluating the needs of culturally and linguistically diverse (CLD) students. One important regulation is that students be assessed in their native language if feasible. In

addition, the American Psychological Association (APA) has postulated several guidelines for psychologists that also reflect the intent of this federal statute (APA, 1993, p.45). The National Association of School Psychologists (NASP) has also endorsed six domains that comprise culturally competent service delivery to individuals in the schools (Rogers, Ingraham, Bursztyn, Cajigas-Segredo, Esquivel, Hess, et al., 1999). Furthermore, the Educational Testing Service (ETS) also has set forth standards to ensure fairness in the design, development, and administration of tests by treating people equally regardless of race, ethnicity, gender, or disability (ETS, 2000).

In order to improve the administration of services to children whose primary language and culture are not those of the majority, school psychologists must be able to accurately identify problem areas with screening instruments that reflect these aforementioned standards. In doing so, school psychologists will assist the public school system to meet the accountability requirements set by IDEA. Given the vast number of children in public schools with culturally and linguistically diverse (CLD) backgrounds, schools are also accountable for addressing the needs of these children. However, in order to do so, it is important that the instruments used meet high standards for both the minority and the mainstream populations. However, only instruments that take these factors into account will provide a valid and accurate representation of a child's ability.

Walton and Vazquez-Nuttall (1992) found that only six screening instruments met their criteria for use with minority preschool populations. This finding clearly illustrates the lack of instruments available for assessing CLD preschool children and accentuates the need for schools to be observant of these issues. Due to the fact that Spanish is the

most prevalent language of children in the schools other than English, there is a need for school psychologists to find instruments that are specifically designed to convey an accurate representation of the abilities of this population. In order to improve the administration of services to children whose primary language and culture are not those of the majority, school psychologists must be able to accurately identify problem areas with culturally sensitive screening instruments. For this reason, not only is it important to evaluate the psychometric properties of screening instruments that are appropriate for young children, but in particular, the psychometric properties of measures that are appropriate for children from CLD backgrounds. Unfortunately, limited information exists regarding screening instruments available for use with students whose native language is Spanish. One screening instrument that has been posited as effective in assessing both English and Spanish-speakers is the Developmental Indicators for the Assessment of Learning – Third Edition (DIAL-3). The instrument directly evaluates children across three domains that assess school readiness: motor, concepts, and language. In the past, children with CLD backgrounds have been misdiagnosed and inappropriately labeled due to a lack of appropriate psychological and/or educational instruments (U. S. D. C., ND, Cal.1970). Researchers in the field of school psychology must be cognizant of the psychometric quality of the instruments used to assess children with CLD backgrounds, so that they are not misdiagnosed. As the field moves to meet the needs of CLD students; researchers, publishers, and practitioners all should be more mindful of the repercussions of having inadequate psychometric properties in the instruments they develop, publish, and/or administer when working with this population.

This study will provide evidence to either support or disconfirm the technical adequacy of using the DIAL-3 to assess both English and Spanish language groups.

Significance of the Study

This study will seek to further examine the psychometric properties of the DIAL-3 for both the English-speaking and Spanish-speaking language groups. One of the objectives of this study is to establish the reliability estimates of the DIAL-3 Spanish version across all six-month age groups in order to ascertain if they approximate the estimates obtained in the English version of the standardization sample. A second objective of this study is to evaluate the discriminant validity in order to provide evidence for separating the three factors as proposed by the test authors. Tests with inadequate discriminant validity do not yield informative scores, consequently, affecting the examiner's ability to provide assistance and/or recommendations for the student. Finally, given that the DIAL-3 is a widely used measure to screen bilingual children, it is important that the construct validity of the instrument be further explored so that practitioners are more confident in their ability to appropriately assess both English-speakers and Spanish-speakers with developmental/academic problems at a young age. The use of psychometrically sound screening instruments for both English-speaking and Spanish-speaking children is critical in order to avoid any potential misdiagnosis.

Research Questions

1. How do the six-month age-group internal consistency reliability estimates for the Spanish version of the DIAL-3 compare to the reported reliability estimates from the English standardization sample?

2. Do the intercorrelations among the DIAL-3 domains in the Spanish-speaking sample provide evidence to support the discriminant validity of the test?
3. Do the models' results for the English standardization sample fit the standardization data for the Spanish version of the DIAL-3? If not, which models better describe the domain(s) evaluated by the test?

CHAPTER II

REVIEW OF RELATED RESEARCH

This chapter will review a series of topics that stress the need for instruments that possess adequate psychometric properties, particularly when evaluating young students from culturally and linguistically diverse (CLD) backgrounds. First, the changing demographics in the United States will be discussed along with the challenges public schools face when addressing the needs of their students. Secondly, critical protective and risk factors will be explored as a way of elucidating the issues young children face. Thirdly, the reader will be made aware of the fact that many young children who come from culturally and linguistically diverse backgrounds do not have the privilege of relying on the same protective factors as their non-minority counterparts. Furthermore, the creation of federal laws to protect the welfare of students who are at-risk, including CLD students will be discussed, while making particular emphasis on the Child Find process delineated by the Individuals with Disabilities Education Act (IDEA). Moreover, the number of young children placed in special education will be reported, making the need for early detection in the schools evident.

An argument will be made for the use of screening instruments as effective and efficient tools for detecting potential disabilities in young children. The importance of selecting instruments that possess adequate psychometric properties and uphold the testing standards of national organizations such as APA, NASP, and ETS will be discussed. Emphasis will be placed on the importance of following the aforementioned guidelines particularly when detecting potential disabilities in culturally and linguistically

diverse young children. The lack of screening instruments that follow the aforementioned criteria to address the needs of young CLD children will be discussed. Finally, the Developmental Indicator for the Assessment of Learning – Third Edition (DIAL-3) will be recommended as an instrument that can be used to address the needs of both English and Spanish-speaking children.

U.S. Demography

Public schools in the U.S. are continually challenged to meet the demands of the children they serve in a manner that is compliant with their governing laws. In addition to this responsibility, parents and government officials require that schools provide them with information that attests to the efficacy of their programs, assuring them that every child is receiving an appropriate education. This task appears to get more difficult with time due to the continual increase in the number of children enrolled in schools. Population estimates indicate that as of 1999 approximately 49.3 million children were enrolled in public school (U.S. Bureau of the Census, 1999). More than half of the student population is currently in elementary school. The student distribution among the different grades for this year was as follows: 5% were in nursery school, 6% were in kindergarten, 59% were in elementary, and 30% were in high school.

Demographic data also shows that public schools are meeting the academic demands of a population that is becoming more ethnically diverse. Estimates indicate that 38% of children enrolled in school are ethnically diverse (U.S. Bureau of the Census, 1999). Of those children who attended schools in 1999, approximately 61% were Caucasian, 17% were Black, 16% were Hispanic, and 5% were Asian/Pacific Islander

(U.S. Bureau of the Census, 1999). Linguistically, the school population has also become diverse with approximately 4.6 million children presently being classified as Limited English Proficient (Kindler, 2002). Reports on language prevalence indicate that Spanish is the most commonly spoken language other than English (Kindler, 2002). Given the aforementioned demographic trends, one can easily understand that being held accountable for the appropriate education of these children is a challenge that can only be met when schools address the needs of their students at an early stage. This is critical given that research has shown that the most effective interventions are those that are implemented early in the child's life (Salisbury & Smith, 1993).

Recent Trends in Hispanic Demographic Characteristics

According to data collected by the Census Bureau on the Current Population Survey (CPS) in 2002, it is apparent that Hispanic population is growing rapidly particularly the Mexican-origin subgroup which makes up 66.9 percent of Hispanics (U.S. Bureau of the Census, 2002). Chapa and Valencia (1993) state that in 1982 there were 15.8 million Hispanics in the United States. The Census Bureau used this number and projected that by the year 2010 the number of Hispanics would double to an approximate 30.8 million. The most recent CPS, however, revealed this projection to have greatly underestimated the Hispanic growth (U.S. Bureau of the Census, 2002). In 2002, there were 37.4 million Latinos representing 13.3 percent of the total U.S. population (U.S. Bureau of the Census, 2002).

Another trend of particular importance for Hispanics is that of educational attainment. Chapa and Valencia (1993) and the U.S. Bureau of the Census (2002) found

that approximately 57 percent of Latinos age 25 had completed high school compared to approximately 80 percent of non-Latinos. Among Hispanics, the Mexican-origin subgroup has the lowest completion rate for both high school and college. According to Chapa and Valencia (1993), only 44 percent of the Mexican-origin subgroup finished high school while only 5 percent completed college. Similarly, the 2002 CPS report found that 50.6 percent of the Mexican-origin subgroup finished high school while only 7.6 percent attained at least a bachelor's degree. In order to change these dim educational attainment trajectories and detect the academic needs of children at an early age, it is important to explore potential factors that will enable students to be successful in school.

Critical Factors that Promote the Academic Success of Young Children

In order to establish what academic need or deficiencies exist in a young child, it is important to first examine the factors that are associated with academically successful children. According to Winquist-Nord and Griffin (1999), the early childhood years are crucial for a child's cognitive and academic development. They believe that it is during these years that children expand their ability to communicate with others and also begin to acquire reading and math skills. Winquist-Nord and Griffin (1999) mention that the success children experience in school is determined by those experiences that take place before the child ever enters a classroom. According to Hart and Risley (1992), a child's learning opportunities depends on the socio-cultural organization of the home as well as the goals parents set through interactions with their child. Hart and Risley (1992) mention that a child's optimal cognitive outcomes are associated with home environments that offer varied opportunities to learn through interaction with caring

adults and age appropriate materials. Hart and Risley (1992) found that the major factors associated with the differences in IQ scores were the extensive amounts of time, attention, and talking that parents invest in their children as well as their interest in what the children have to say. Moreover, Rebello-Britto and Brooks-Gun (2001) explored the influence of the home environment on children's literacy development as well as school achievement skills. According to the aforementioned authors, numerous experiences that foster the active engagement with meaningful forms of reading, writing, and spoken language as well as a supportive environment that encourages learning all contribute to children becoming readers. Rebello-Britto and Brooks-Gun's (2001) review of the literature found that better vocabulary skill in early and middle childhood is evidenced when there was early exposure to language in the home. The authors also mention the importance of decontextualized language (i.e. the ability to talk about ideas and non-present objects) on emergent literacy and believe that this is a critical ingredient to school success. According to Rebello-Britto and Brooks-Gun (2001), the social-emotional climate is another important influencing factor in a student's academic success. The authors indicate that a parent's supportive presence and warmth are associated with a child's language ability.

Risk Factors Associated with Low Academic Performance

Unfortunately, many of the children, especially those from culturally and linguistically diverse backgrounds, do not have the privilege of relying on the aforementioned learning opportunities and/or resources. Therefore, it is important to also become familiar with the risk factors that hinder the academic success of children.

Research has demonstrated that there are certain social and environmental factors that negatively impact the academic performance of young children (Zill, 1996; Portes & MacLeod, 1996; McLanahan & Sandefur, 1994). Some of those variables include poverty, maternal depression, the educational attainment of parents, and living in a single parent household. Several studies have indicated that poverty is linked to low achievement test scores, grade repetition, problem behaviors that result in suspension/expulsion, and dropping out of school; all factors that can be translated into poor educational outcomes for children (Zill et al., 1995; McLanahan & Sandefeur, 1994). According to Snow and Paez (2004), language development in young children can be associated with their later school success in both reading and acquiring content knowledge. Several studies have established that children from low socio-economic status tend to have a diminished vocabulary size as compared to their more affluent counterparts, which has also shown to be connected to the vocabulary used by the parent (NRC, 1998; Hart and Risley, 1992, 1995; Davidson, 1993). According to a study conducted by Smith, Brooks-Gunn, and Klebanov (1997), differences in the home learning environments of high SES and low SES can explain as much as half of the achievement gap in achievement tests in preschool children.

Moreover, in some instances, poverty also can be a catalyst for depression in mothers (Edin & Lein, 1997). Bettes (1988) found that maternal depression appears to be associated with linguistic as well as emotional development. Furthermore, the National Research Council (2000) reports that maternal depression has been linked to a reduction in the quality of parenting as well as to problems in the emotional relationship between

parents and their children. These children are at a higher risk for emotional and behavioral problems which have been associated with difficulties in school (Campbell & Ramey, 1995). With regard to the educational attainment of parents, research suggests that this factor alone can influence the level of education and income their children achieve as adults (Hoff, Laursen, & Tardiff, 2002). The number of parents living in a household also impacts the academic success of the child. According to McLanahan and Sandefur (1994), those children who live with two biological parents do better in an array of domains, including school, when compared to children who live with only one parent or no parents. In order to protect the welfare of children with academic needs and who oftentimes evidence one or more of the aforementioned factors, several laws were enacted to provide them with opportunities for success in school.

Federal Statutes

Children with disabilities may be protected by three federal statutes: Section 504 of the Rehabilitation Act of 1973 (Section 504), the Americans with Disabilities Act of 1990 (ADA), and the Individuals with Disabilities Education Act, Part C (IDEA). IDEA is the primary federal law entitling children with disabilities to appropriate educational services. Part C of IDEA requires that states provide children between the ages of 3 and 21, who have disabilities that affect their educational performance, with a free appropriate public education (FAPE) in the least restrictive environment (LRE). Section 504 of the Rehabilitation Act prohibits discrimination on the basis of disability by any program that receives federal funding, including public school programs. Children are protected under Section 504 if they have a physical or mental impairment which

substantially limits one or more major life activity, if they have a record of having such an impairment, or if they are regarded as having such an impairment. The ADA prohibits discrimination on the basis of disability by public entities, including public schools. Children are protected by provisions of Title II of the ADA if they have a physical or mental impairment which substantially limits one or more major life activity, if they have a record of having such an impairment, or if they are regarded as having such an impairment.

Early Identification and the Child Find Process

One critical component in the school accountability process is its Child Find process. According to the Individuals with Disabilities Education Act (IDEA), schools are responsible for identifying, locating, and evaluating all preschool children who may evidence a potential disability and subsequently, are in need of early intervention or special education services (P.L. 105-17). According to Thurlow (1992), prior to 1975 a large number of children were entering school with difficulties that impacted their ability to learn as well as children that had academic problems that could have been addressed when they were younger. Accountability for these children can easily become a problem if preventative steps are not implemented to identify problems early to be in compliance with the aforementioned federal statute. Preventative steps can give schools the opportunity to address the needs of its students before systemic inefficiencies can negatively impact student academic outcomes. Consider the consequences for a young child with a developmental disability whose academic trajectory is negatively impacted by the lack of early identification; a situation that could have been easily avoided if

schools had a screening program in place that would have prevented this from happening. By taking preventative steps to identify children early, schools will be better prepared to meet the public's demand for accountability and to safeguard their students' education.

Early Childhood and Special Education

Since one of the goals of IDEA is to determine the academic needs of children at an early stage, identification of these students should take place before or at the beginning of their academic career in the public school system approximately during ages 3 to 5 or 6 years. In doing so, the learning potential of these students can be maximized since they will be provided with the necessary attention and academic support to be successful in school. Approximately 1 million children between the ages of 3 to 6 years are currently being served under IDEA (U.S. Department of Education, 2003). The total number of children receiving special education services in the 2000-2001 school year by age group is as follows: 127,281 children age 3 years, 208,677 children age 4 years, 263,720 children age 5 years, and 332,016 children age 6 years (U.S. Department of Education, 2003). Information for the three most frequently prescribed disability categories to children in this age span are presented in this section. An examination of the information presented in the 24th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act (U.S. Department of Education, 2003) revealed a clear pattern with regard to the most frequently used disability categories for ages 3 to 6 years. In these age groups, the most prevalent disability category was Speech and Language Impairments. Of those children who received special education services,

Speech or Language Impairments constituted 46.3% of the age 3 group, 52.1% of the age 4 group, 61.9% of the age 5 group, and 64.2% of the age 6 group. With the exception of the age 6 group, the second most frequent disability category was Developmental Delay, making up 33.4% of the age 3 group, 29.9% of the age 4 group, and 17% of the age 5 group. The second most frequent disability category for those children who were age 6 and served under IDEA was Specific Learning Disabilities, comprising 10.6% of this group. The difference in prevalence between the younger age groups and the 6 year old group can be attributed to the fact that more academic demands are made of older students, making this disability more salient. The third most frequent disability category among all age groups was Mental Retardation which classifies 4.1% of children in the age 3 group, 3.7% of the age 4 group, 4.8% of the age 5 group, and 6% of the age 6 group. As student enrollment in public schools increases, and given the large number of children served under IDEA during the 2000-2001 academic year, schools should expect a similar, if not greater, number of children to be deemed eligible to receive special education services in the future. Moreover, since schools are being held accountable for these children, it is important that these public institutions implement systems that can detect and address the needs of these children early on in their lives.

Early Intervention

According the National Research Council (NRC) (2000), interventions that begin early and are maintained appear to be more beneficial for young children. Ramey and Ramey (1999) suggest that high-risk children fall behind as early as the second year of life if no intervention is implemented. Studies have shown sustained benefits in cognitive,

social and achievement areas when programs are implemented to offset these deleterious effects (Campbell & Ramey, 1995; Lazar, Darlington, Murray, Royce & Snipper, 1982; Schweinhart & Weikart, 1997). These benefits are often enough to keep some children with mild disabilities from receiving special supports from the school system (NRC, 2002). Moreover, a study conducted by Reynolds, Temple, Robertson, and Mann (2001) demonstrated that preschool participation was associated with significantly lower rates of special education placement. One early intervention program that was developed to narrow the gap between low SES children and their more affluent counterparts was the Head Start program (U.S. General Accounting Office, 1998). Head Start is a federally funded program that provides educational experiences, improved nutrition, increased parental involvement, and access to health and social services (U.S. General Accounting Office, 1998). Studies that have examined the trajectory of Head Start children have indicated that those children that began to learn about print, sounds, and writing during the early years were more likely to be reading successfully in elementary school (Lonigan, Burgess, & Anthony, 2000; Storch & Whitehurst, 2002; Whitehurst & Fischel, 2000). Overall, it appears that by developing a system of early interventions through specialized programs such as Head Start, schools can begin to address the needs of young children who are at-risk and change their trajectories to ones with a greater opportunity for academic success.

The Role of School Psychologists in Early Intervention

To ease the onus placed on the schools, it is important that school psychologists and other diagnostic personnel remain vigilant regarding the most efficient and cost

effective means to identify problems early. Unfortunately, due to the large numbers of children with academic needs in addition to a limited school budget, schools cannot afford to evaluate every child comprehensively when a child's competencies are in question. Therefore, an alternative way in which school psychologists can accomplish this task is through the use of screening instruments. According to Meisels (1985), screening instruments are "a brief assessment procedure designed to identify children who, because of risk of a possible learning problem or handicapping condition, should proceed to a more intensive level of diagnostic assessment" (p.1). However, assessing young children can be difficult given that they have different rates of maturation, different experiential backgrounds, as well as different cultural and linguistic backgrounds (Romero, 1999). Since these factors impact the performance of the child during an evaluation, it is important that the instruments used meet the highest standards so that their influence is minimized. The use of quality instruments will ensure that the assessments conducted accurately identify children with academic needs at an early stage, giving schools the necessary information to design interventions that meet those needs.

IDEA, Ethics, and Testing Standards in working with Culturally and Linguistically Diverse Populations

Individuals with Disabilities Act (IDEA)

IDEA clearly stipulates several procedures that must be addressed when evaluating the needs of culturally and linguistically diverse (CLD) students. IDEA indicates the following: (1) standardized tests must be administered by trained and knowledgeable personnel; (2) standardized tests must be validated for the task for which they are used

(20 U.S.C. § 1414(b)(3)(B)); (3) tests that are not racially or culturally discriminatory toward the child being evaluated must be used (20 U.S.C.S. § 1414(b)(3)(A)(ii)); and (4) students must be tested in their native language or mode of communication (34 C.F.R. § 300.532(a)(2)).

Ethical Guidelines for Practitioners

Several organizations recognized the importance of addressing the factors that impact the assessment of culturally and linguistically diverse learners and created ethical guidelines and standards that fall in suit with the principles behind the enactment of IDEA. The American Psychological Association (APA) is one of the organizations that postulated guidelines for psychologists. The main recommendations by APA (1993) for providing psychological services to ethnically, linguistically, and culturally diverse populations are listed below:

1. Psychologists educate their clients to the processes of psychological intervention, such as goals and expectations; the scope and, where appropriate, legal limits of confidentiality; and the psychologists' orientations.
2. Psychologists are cognizant of relevant research and practice issues as related to the population being served.
3. Psychologists recognize ethnicity and culture as significant parameters in understanding psychological processes.
4. Psychologists respect the roles of family members and community structures, hierarchies, values, and beliefs within the client's culture.
5. Psychologists respect clients' religious and/or spiritual beliefs and values, including attributions and taboos, since they affect world view, psychosocial functioning, and expressions of distress.
6. Psychologists interact in the language requested by the client and, if this is not feasible, make an appropriate referral.
7. Psychologists consider the impact of adverse social, environmental, and political factors in assessing problems and designing interventions.
8. Psychologists attend to as well as work to eliminate biases, prejudices, and discriminatory practices.
9. Psychologists working with culturally diverse populations should document culturally and socio-politically relevant factors in the records (p.45).

In addition to the standards stipulated by APA, the National Association of School Psychologists (NASP) has also endorsed six domains that comprise culturally competent service delivery to individuals in the schools (Rogers, Ingraham, Bursztyn, Cajigas-Segredo, Esquivel, Hess, et al., 1999). The six domains are as follows:

I. Legal and Ethical Issues
   a. Knowledge of local, state, and federal laws and regulations, and court rulings pertaining to culturally and linguistically diverse children and youth
   b. Advocate for public policy and educational laws that eliminate disparities in services to diverse children and youth
   c. Understanding of ethical standards as they relate to delivering services to culturally and linguistically diverse children and youth
   d. Ability to recognize the limits of our own multicultural competence
   e. Seek educational, consultative, and training experiences to improve multicultural knowledge

II. School Culture, Educational Policy, and Institutional Advocacy
   a. Knowledge of and advocacy for aspects of organizational culture and values that promote achievement and mental health and reduce risk of inappropriate services for culturally and linguistically diverse (CLD) students
   b. Examine individual referrals within the context of institutional and systemic patterns (e.g. cultural misinformation, racism, cultural differences) affecting CLD learners and provide leadership in seeking and implementing systemic interventions.
   c. Skills in educating the school and community about issues which affect the learning, development, and well-being of children from CLD backgrounds

III. Psychoeducational Assessment
   a. Knowledge of and skills in assessing CLD students, including consideration of variables such as environment, social issues, language development, second language acquisition, acculturation, educational history, quality of educational program, SES and racism
   b. Understanding that normed tests may not be a valid measure for English Language Learners (ELLs) due to inappropriateness of norms, scores reflecting English proficiency, product as opposed to process orientation, fairness of content, and differences in educational background, acculturation, and economic situation

IV. Academic, Therapeutic, and Consultative Interventions
   a. Skills in multicultural counseling and cross-cultural consultation

b. Knowledge of multicultural education, ELL programs, and school culture/culture of staff and students

V. Working with Interpreters
   a. Knowledge of recommended systemic practices, including guidelines from professional organizations and national and state policies, and plans for hiring, training, and managing interpreters
   b. Knowledge of recommended practices for interpreters translating for parent conferences, including using school personnel and community members as interpreters (never children or family members)

VI. Research
   a. Knowledge of research related to culture and language issues and ability to conduct research using qualitative and quantitative methods that is sensitive to cross-cultural issues
   b. Knowledge of how own value system and identity impacts the formulation of research questions, the selection of research methods, the development of research design, and interpretations made of the results
   c. Skills in eliminating bias when conducting research
   d. Awareness of Emic-Etic distinctions (Emic: behaviors or views that are common to an ethnic or minority group; Etic: aspects of human functioning that are more universal to peoples across cultures) (p.243)

Testing Standards

Since schools and school psychologists must select instruments with high standards, it is important to provide an overview of what constitutes a quality instrument. Knowledge of these criteria will assist the school psychologists obtain valid information regarding the preschool child. Moreover, knowledge of these criteria will also serve to help schools in meeting the demand for accountability, since any screening instrument that incorporates these standards will yield the best information regarding the academic needs of children.

The Educational Testing Service (ETS) also has set forth standards to ensure fairness in the design, development, and administration of tests by treating people equally regardless of race, ethnicity, gender, or disability (ETS, 2000). There are eight standards that address fairness in testing:

1. Address fairness in the design, development, administration, and use of the product or service, and document what was done. Provide a plan for addressing fairness for a product or service under development or facing major revision.
2. Obtain and document judgmental and, when possible, empirical evaluations of fairness for studied groups. As appropriate, ensure that various groups are represented. Ensure that symbols, language, and content that are generally regarded as sexist, racist, or offensive are eliminated, except when necessary to meet the purpose of the assessment, product, or service.
3. Provide impartial access to products and services. For assessments, provide impartial registration, administration, and reporting of assessment results.
4. When a construct can be measured in different ways that are reasonably equally valid, reliable, practical, and affordable, consider available evidence of group differences in assessment results in determining how to measure the construct.
5. Provide appropriate and reasonable accommodations for people with disabilities, in accordance with applicable laws and client's policies.
6. If there is evidence of the comparability of scores between standard and modified assessments or administrations, do not flag scores as nonstandard. If there is insufficient evidence of the comparability of scores, indicate the nature of the accommodation when possible.
7. Consider the needs of nonnative speakers of English in the development and use of products or services. For assessments, reduce threats to validity that may arise from language differences.
8. For research studies, perform separate analyses for studied groups when the information is relevant and sample sizes are sufficient. Obtain informed consent for participation of human subjects, and avoid negative consequences of participation for members of all groups. (p.18)

The Code of Fair Testing Practices in Education lists the following criteria with respect to the development and selection of appropriate tests (Joint Committee on Testing Practices, 2003):

1. (*Test authors should*) provide evidence of what the test measures, the recommended uses, the intended test takers, and the strengths and limitations of the test, including the level of precision of test scores.
2. Describe how the content and skills to be tested were selected and how the tests were developed.
3. Communicate information about a test's characteristics at a level of detail appropriate to the intended test users.
4. Provide guidance on the levels of skills, knowledge and training necessary for appropriate review, selection, and administration of tests.

5. Provide evidence that the technical quality, including reliability and validity, of the test meets its intended purposes.
6. Provide to qualified test users representative samples of test questions or practice tests, directions, answer sheets, manuals, and score reports.
7. Avoid potentially offensive content or language when developing test questions and related materials.
8. Make appropriately modified forms of tests or administration procedures available for test takers with disabilities who need special accommodations.
9. Obtain and provide evidence on the performance of test takers of diverse subgroups, making significant efforts to obtain sample sizes that are adequate for subgroup analyses. Evaluate the evidence to ensure that differences in performance are related to the skills being assessed (p.3).

All in all, the guidelines and standards postulated by the Joint Committee on Testing Practices (standard 9), the American Psychological Association (standard 8), the National Association of School Psychologists (domain III(a) and III(b)), and the Education and Testing Services (standard 7), have emphasized the importance of addressing the factors that are involved in assessing children with culturally and linguistically diverse backgrounds. In order to improve the administration of services to children whose primary language and culture are not those of the majority, school psychologists must be able to accurately identify problem areas with screening instruments that reflect these aforementioned standards. In doing so, school psychologists will assist the public school system to meet the accountability requirements set by IDEA.

Validation Procedures for the Development of Culturally Sensitive Instruments

Research suggests that there are four types of equivalents that must be considered in order to establish the cross-cultural validity of an instrument. The first type of equivalents is called linguistic equivalent and refers to having an accurate translation of

behavioral descriptors. If this type of equivalence is not established than behavioral raters may not have a common understanding of the characteristic being rated. In order to satisfy linguistic equivalence in an instrument back translation is recommended. To do this, a word is first translated into a second language and then retranslated back into the original language. If the retranslation does not yield the same word than as the translation did, than the translation did not have an equivalent literal meaning (Marsella & Kameoka, 1989). A second characteristic of a cross-culturally validated instrument is that of conceptual equivalence and refers to the similarities found in the meaning of the constructs used in assessment. Different cultures can interpret a construct in different ways, for example the term "dependency" has a negative connotation in Western societies yet it has a positive connotation in Japanese society (Marsella & Kameoka, 1989). Moreover, a study conducted by Reid, Maag, and Vasa (1994) revealed that there are clear differences in the perceptions of certain disabilities across European, British, and American professionals. The third characteristic of a cross-culturally valid instrument is scale equivalence. In order to meet this criterion there must be a common understanding of how the scale is used and that the raters share a common metric. In an attempt to quantify an opinion, behavior rating scales usually use Likert Scales using descriptors such as "Not at All," "Just a Little," "Pretty Much," and "Very Much" (Marsella & Kameoka,1989). Different cultures may interpret the frequency, intensity, and duration of the behavior corresponding to each of the aforementioned descriptors resulting in interrater differences within a culture (Ross & Ross 1982). The last characteristic constituting a culturally valid instrument is that of normative equivalence

which implies that the standards developed for one culture are appropriate for another (Marsella & Kameoka (1989). According to Reid (1995) in order for instruments to be comparable across cultures, distribution should be similar, population means should be equal, and the symptomatology of a disorder should take place at the same base rate and intensity across cultures. Objective instruments can be very useful when trying to understand a child's academic and/or behavioral needs, however, practitioners should be aware that the results obtained when using an instrument cross-culturally may be inappropriate given that it is very rare to find instruments that meet all four criteria.

Culturally Sensitive Preschool Screening Instruments

Since Spanish is the most prevalent language of children in the schools other than English, there is a need for school psychologists to find screening instruments that are specifically designed to convey an accurate representation of the abilities of this population. For this reason, not only is it important to evaluate the psychometric properties of screening instruments that are appropriate for young children but also the psychometric properties of measures that are appropriate for children from CLD backgrounds. Walton and Vazquez-Nuttall (1992) found that only six screening instruments met their criteria for use with minority populations: 1) The Cooperative Pre-School Inventory, Revised-Spanish Translation, 2) Dallas Pre-School Screening Test-Spanish Translation, 3) The Denver Developmental Screening Test-Spanish Translation, 4) Early Screening Inventory-Spanish Translation, 5) Preschool Screening System-Translations in Various Languages and 6) The Developmental Indicator for the Assessment of Learning-Revised. The main reason being that their standardization

samples included minority populations, had translations or adaptations in several languages, or had specific reliability and validity in studies conducted using minority groups (Walton & Vazquez-Nuttall, 1992). This finding clearly illustrates the lack of instruments available for assessing young CLD children and accentuates the need for schools to be observant of these issues. This is especially true since the norms of the aforementioned instruments appear to be outdated. For example, the oldest norms correspond to the Cooperative Pre-School Inventory, Revised-Spanish Translation in 1970, while the most recent norms correspond to the Early Screening Inventory-Spanish Translation in 1988. Practitioners should be aware that the norms for these instruments may not be representative of the Spanish-speaking population of today and should be used with caution. According to Lidz (2003), examiners must understand a test's strengths and weaknesses so that they can be used appropriately. Given the information discussed in this section regarding culturally sensitive screeners, it is reasonable to conclude that limited information exists regarding screening instruments available for use with students whose native language is Spanish. This study will expand the available information regarding the psychometric properties of screening instruments used in early childhood. This information will be useful to both the mainstream culture and to minority cultures whose native language is Spanish. Specifically, the study will address the construct validity of both the English version and the Spanish version of the Developmental Indicators for the Assessment of Learning - Third Edition (DIAL-3).

Developmental Indicators for the Assessment of Learning – Third Edition (DIAL-3)

Description

One screening instrument that has been posited as effective in assessing both English and Spanish-speakers is the Developmental Indicators for the Assessment of Learning – Third Edition (DIAL-3). This instrument was developed by Carol Mardell-Czudnowski and Dorothea S. Goldenberg as an improvement over its predecessors, the DIAL and the DIAL-R, and was published in 1998 by American Guidance Services. The DIAL-3 can be administered in either English or Spanish to children ranging from ages 3 years 0 months to 6 years 11 months. The instrument directly evaluates children across three domains that assess school readiness: motor, concepts, and language. This measure has a total of twenty-one items separated into seven items per domain. Depending on the age group or on task performance, domains may contain only six items. For example, children age 3 years do not receive the writing task in the motor domain (item 7) since writing is a skill that has not been developed by that time. Each item yields a specific raw score based on the child's performance and is immediately converted into a scaled score ranging from 0 to 4. Each obtained scaled score has a corresponding age level expectation. A child obtaining a scaled score of 0 indicates that the child's performance is what one would have expected for a child that is under 3 years old. A child obtaining a scaled score of 1 indicates that the child's performance is what one would have expected for a child that is 3 years old. A child obtaining a scaled score of 2 indicates that the child's performance is what we would have expected for a child that is 4 years old. A child obtaining a scaled score of 3 indicates that the child's performance is what we

would have expected for a child that is 5 years old. A child obtaining a scaled score of 4 indicates that the child's performance is what one would have expected for a child that is 6 years old. These scaled scores are then recorded onto the DIAL-3 protocol. The sums of the scaled scores are then converted into percentile ranks and subsequently into standard scores. The sum of the scaled scores for each domain are also added together to yield a Total score, which then undergoes the aforementioned score transformation process. Once standard scores are obtained for the three domains and the total score has been calculated, the examiner proceeds to select cutoff scores. These cutoff scores help the examiner evaluate whether or not the child is exhibiting a potential delay in any of the four scores. The tables presented in the DIAL-3 manual contain the cutoff scores that correspond to 5 different confidence intervals: 84%, 90%, 93%, 95%, and 98%. Examiners are encouraged to select the confidence intervals that best suit their needs.

Internal Consistency Reliability

The internal consistency of the DIAL-3 was estimated by computing Cronbach's alpha for each age group in the sample per domain. The manual provides reliability estimates for each six-month age group in the English standardization sample. The reliabilities in the motor domain ranged from an $\alpha = .45$ to an $\alpha = .74$. Reliabilities in the concepts domain ranged from an $\alpha = .65$ to an $\alpha = .88$. The language domain had reliabilities ranging from an $\alpha = .66$ to an $\alpha = .80$. Lastly, the reliabilities for the total scores ranged from an $\alpha = .77$ to an $\alpha = .90$. The median reliabilities for the motor, concepts, language, and total scores in the English version were $\alpha = .66$, $\alpha = .84$, $\alpha = .77$,

and α = .87, respectively. Overall, these reliability estimates provide evidence supporting a consistency in the scores within each domain.

Reliability estimates for the Spanish standardization sample were not reported in the DIAL-3 manual; however, they were reported in a related article by Chen, Wang, Czudnowski, Goldenberg, and Elliot (2000). The Cronbach's Alphas for the Spanish version are not as explicit as those in the English version since the age groups were collapsed into twelve-month groups and not six-month groups. Chen et al. (2000) reported that the reliabilities in the motor domain ranged from an α = .39 to an α = .67. Reliabilities in the concepts domain ranged from an α = .69 to an α = .83. They also found that the language domain had reliabilities ranging from an α = .74 to an α = .77. Lastly, the reliabilities for the total scores ranged from an α = .78 to an α = .87. Chen et al. (2000) also reported the following median reliabilities for the Spanish version: α = .64 for the motor domain, α = .81 for the concepts domain, α = .74 for the language domain, and α = .86 for the total score. The reliability coefficients reported are similar to those in the English standardization sample and also provide evidence supporting the consistency of the scores within each domain. Nonetheless, there is not enough information provided by Mardell-Czudnowski and Goldenberg (1998) to determine if the same reliabilities will hold once the sample is divided into six-month age groups. Collapsing the reliability estimates into twelve month groups may mask significant reliability changes among the age groups. This aspect is important to examine since considerable changes in the estimates could make the instrument adequate for certain age groups but not for others.

Test-Retest Reliability

Test-Retest reliabilities were computed for both the English and the Spanish standardization samples. The authors report a median day interval of 28 days for both samples. The sample for the English version included 158 children while the Spanish version only included 37 children. The test-retest reliabilities in both samples were corrected for norm group variability (SD=15) using Guilford's formula, based on the standard deviation obtained on the first testing period (Guilford, 1954). The Motor domain had a test-retest reliability of .66 for the English sample and .53 for the Spanish sample (Chen et al., 2000). The Concepts domain had a test-retest reliability of .80 for the English sample and .88 for the Spanish sample. The Language domain had a test-retest reliability of .82 for the English sample and .81 for the Spanish sample. Finally, the Total score had a test-retest reliability of .86 for the English sample and .87 for the Spanish sample (Chen et al., 2000). Overall, the reliability estimates appear to be fairly similar between samples and indicate a small impact of time on the scores. These estimates may only be moderately stable given the small sample sizes used to determine these coefficients.

Discriminant Validity

The discriminant validity of the English version of the instrument was established by examining the intercorrelation among the three factors that constitute the DIAL-3. The motor and concepts domain had an intercorrelation of r = .50. The motor and the language domain had a correlation of r = .41. Finally, the correlation between the concepts and language domain was r = .65. The correlation between the concepts domain

and the language domain of the test is moderately high with both domains sharing 42% of the variance. However, the reported intercorrelations apply only to the standardization sample for the English version of the test. The discriminant validity of the test was not evaluated in the sample for the Spanish version. This may be a potential threat to the discriminant validity of the test since the correlation indicates a higher degree of similarity among the factors than is desired.

Concurrent Validity

The concurrent validity of the instrument was assessed by correlating the DIAL-3 scores from the English version with other screening instruments, measures of intelligence, and language measures. Six measures that were used to evaluate the concurrent validity of the test were: the Early Screening Profiles (ESP), the Battelle Developmental Inventory Screening Test (BDIST), the Bracken Basic Concept Scale (Screening Test - BBCS), the Brigance Preschool Screen, the Differential Abilities Scales (DAS), and the Peabody Picture Vocabulary Test – Third Edition (PPVT-III).

The Early Screening Profiles test was given to 76 children ranging from age 3 years 8 months to age 5 years 8 months who also were represented in the standardization sample. This test yields 9 subscale scores: the Cognitive/Language Profile, the Cognitive subscale, the Language subscale, the Self-Help/Social Profile, the Verbal Concepts subscale, the Visual Discrimination subscale, the Logical Relations subscale, and the Basic School Skills subscale. However, only the validity coefficients for the subscales that are most closely associated with the DIAL-3 will be reported. The correlation between the Language subscale and the DIAL-3 language domain was $r = .38$. The Motor

Profile subscale and the DIAL-3 Motor domain had a correlation of r = .42. Lastly, the correlation between the Verbal Concepts subscale and the DIAL-3 Concepts domain was r = .45.

The Batelle Developmental Inventory Screening Test (BDIST) was given to 71 children ranging from age 3 years 2 months to age 5 years 10 months who also were represented in the standardization sample. The BDIST yields six subscale scores and a total score. The correlation for the BDIST Total score and the DIAL-3 Total score was r = .79.

The Brigance Preschool Screen was given to 28 children ranging from age 3 years 7 months to age 4 years 11 months who also were represented in the standardization sample. The Brigance only yields a Total score. The correlation between the Brigance Total score and the DIAL-3 Total score was r = .54.

The Differential Abilities Scales (DAS) was given to 50 children ranging from age 3 years 7 months to age 5 years 10 months who also were represented in the standardization sample. The DAS yields 10 subscale scores and 3 composite scores. The correlation between the General Conceptual Ability score and the DIAL-3 Total score was r = .75.

The Peabody Picture Vocabulary Test – Third Edition (PPVT-III) was given to 49 children ranging from age 3 years 6 months to age 4 years 11 months. The test was also given to a second sample of children ranging from age 5 years 0 months to 6 years 11 months. All children also were represented in the standardization sample. The PPVT-III yields a total language score. The language score from the PPVT-III was correlated with

the language domain of the DIAL-3 yielding a coefficient of r = .63 for the first sample and a coefficient of r = .45 for the second sample.

As in the case with the discriminant validity of the instrument, the concurrent validity of the instrument was established using the standardization sample for the English version only. No data is available on the concurrent validity of the instrument for the Spanish-speaking sample.

Given the various technical limitations of the DIAL-3 with regard to the Spanish-speaking population, it is important to expand the work of the DIAL-3 authors to include more detailed information regarding its reliability and validity. In particular, the following section will enumerate the questions that will be addressed in this study.

Research Questions

1. How do the six-month age-group internal consistency reliability estimates for the Spanish version of the DIAL-3 compare to the reported reliability estimates from the English standardization sample?
2. Do the intercorrelations among the DIAL-3 domains in the Spanish-speaking sample provide evidence to support the discriminant validity of the test?
3. Do the models' results for the English standardization sample fit the standardization data for the Spanish version of the DIAL-3? If not, which models better describe the domain(s) evaluated by the test?

Significance of the Study

Given the research design for this study, several significant outcomes are anticipated. First, the results to research question one will reveal if the DIAL-3 has

adequate reliability estimates across all six-month age groups that approximate the estimates obtained in the English version of the standardization sample. This will be one of the pieces that determine if the instrument is appropriate for use with the Spanish-speaking population. Secondly, evaluating the discriminant validity of the test will provide evidence for separating the three factors proposed by the test authors. Tests with inadequate discriminant validity do not yield informative scores, consequently, affecting the examiner's ability to provide assistance and/or recommendations for the student. Thirdly, the confirmatory factor analyses for the different age groups will yield empirical evidence to support or reject the existence of three domains within the DIAL-3. Results of the confirmatory factor analysis may be different from what was suggested by the authors and could invalidate the use of the DIAL-3 for one or more age groups. Cognitive components like Concepts and Language are very hard to differentiate in young children. As children grow older, these cognitive components may emerge, allowing psychologists to measure them. Given that the DIAL-3 is a widely used measure to screen bilingual children, it is important that the technical properties be further explored so that practitioners are more confident in their ability to appropriately refer both English-speakers and Spanish-speakers with developmental/academic problems at a young age. The use of psychometrically sound screening instruments for both English-speaking and Spanish-speaking children is critical in order to avoid any potential misdiagnosis.

CHAPTER III

METHODOLOGY

This chapter will describe the demographics of the sample and discuss the psychometric properties reported in the DIAL-3 technical manual. Moreover, this chapter will present a discussion of the procedures that were followed, the instruments used for the study, the proposed research questions as well as the statistical analyses used to answer each research question.

Participants

The present study will be conducted using the data from the standardization samples of both the English and the Spanish versions of the DIAL-3. The following sections reflect the available participant information for these abovementioned standardization samples.

English Standardization Sample

The standardization sample included 1,560 children for the English version of the DIAL-3. The sample for the English version of the test was obtained in such a way that it closely matched the 1994 Current Population Survey (CPS) for March. The Current Population survey is a monthly survey of approximately 50,000 households conducted jointly by the Labor Statistics Bureau and the Census Bureau that yields a number of variables regarding demography and employment. The survey uses multistage sampling and is considered to be a representative measure of households in the United States. The age-groups in the DIAL-3 were stratified using the following four variables: gender, race/ethnicity, geographic region, and parental education. The number of children in each

stratum was compared to the strata in the CPS to assess the representativeness of the standardization sample. Those strata that did not approximate the numbers found in the CPS were given sampling weights to control for this problem. Consequently, the standardization sample for the English version of the DIAL-3 appears to be representative of the U.S demography.

Chronological Age. The children participating in this study were divided into six-month age groups starting at age 3 years 0 months and ending at age 6 years 11 months. The number of children in the 3-0 to 3-5 age group was 110. The number of children in the 3-6 to 3-11 age group was 197. The number of children in the 4-0 to 4-5 age group was 234. The number of children in the 4-6 to 4-11 age group was 297. The number of children in the 5-0 to 5-5 age group was 274. The number of children in the 5-6 to 5-11 age group was 165. The number of children in the 6-0 to 6-5 age group was 162. The number of children in the 6-6 to 6-11 age group was 121. The total sample used to develop the English version of the DIAL-3 was comprised of 1,560 children.

Gender. The number of males and females in the sample was approximately equal. Females comprised 46.9% (n = 732) of the total sample, while males comprised 53.1% (n = 828) of the total sample.

Geographic Region. Participants were placed in one of four geographic regions: Northeast, North Central, South, and West. The Northeast group was comprised of children in the following states: Maine, Massachusetts, New Hampshire, New York, and Pennsylvania. The North Central group was comprised of children in the following states: Illinois, Indiana, Kansas, Michigan, Missouri, North Dakota, Ohio, South Dakota, and

Wisconsin. The South group was comprised of children in the following states: Alabama, Arkansas, Florida, Georgia, Kentucky, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and the District of Columbia. The West group was comprised of children in the following states: Arizona, California, Colorado, Hawaii, Idaho, Montana, Oregon, Washington, and Wyoming. One hundred eighty five children were in the Northeast group, four hundred sixty five children were in the North Central group, six hundred ninety five children were in the South group, and two hundred fifteen children were in the West group.

Parental Education Level. The majority of children had parents with one to three years of college (n = 552, 35.4%), followed by parents who were high school graduates (n = 539, 34.6%), parents with a grade 11 education or less (n = 248, 15.9%), and parents with four or more years of college (n = 221, 14.2%).

Spanish Standardization Sample

The standardization sample for the Spanish version of the test was not compared to the CPS or to any other Census report to assess the representativeness of the sample, unlike the English version of the test. The data collected for this version of the test consisted of 536 monolingual Spanish-speaking children in the U.S., 26 children from Panama, and 43 children from Puerto Rico for a total sample size of 605 children. Nevertheless, the age-groups in the Spanish version of the DIAL-3 were stratified into the same four variables as in the English version. Unfortunately, given the limited information presented in the manual, it is difficult to determine whether the sample

reflects the Spanish-speaking population in the United States or is only a subset of this population.

Chronological Age. The children participating in this study were divided into six-month age groups starting at age 3 years 0 months and ending at age 6 years 11 months. The number of children in the 3-0 to 3-5 age group was 25. The number of children in the 3-6 to 3-11 age group was 51. The number of children in the 4-0 to 4-5 age group was 81. The number of children in the 4-6 to 4-11 age group was 176. The number of children in the 5-0 to 5-5 age group was 120. The number of children in the 5-6 to 5-11 age group was 76. The number of children in the 6-0 to 6-5 age group was 45. The number of children in the 6-6 to 6-11 age group was 31. The total sample used to develop the Spanish version of the test was comprised of 605 children.

Gender. The number of males and females in the sample was approximately equal. Females comprised 46.4% (n = 281) of the total sample, while males comprised 53.6% (n = 324) of the total sample.

Geographic Region. Participants were placed in the same four geographic regions as in the English standardization sample, except that the children who were tested in Puerto Rico and Panama were added on to the South group. Ten children were in the Northeast group, twenty eight children were in the North Central group, one hundred eight children were in the South group, and four hundred fifty nine children were in the West group. However, the South group only included thirty nine children who were tested in the United States. Sixty-nine children were tested in either a U.S. territory or another country.

Parental Education Level. The majority of children had parents with a grade 11 education or less (n = 424, 70.1%), followed by parents with one to three years of college (n = 72, 11.9%), parents who were high school graduates (n=70, 11.6%), and parents with four or more years of college (n = 39, 6.4%). The DIAL-3 manual did not specify the number of parents living in the U.S., the number of parents living in U.S. territories, or the number of parents living in other countries.

Procedures

A request to purchase the standardization database for the English-speaking sample and the Spanish-speaking sample was sent to American Guidance Services, Incorporated. The publisher approved the purchase of the DIAL-3 standardization data and the data was received in ASCII format. All information in the database was coded in order to protect the confidentiality of the participants. Consent for participation in the study was obtained from the parents by the testing site coordinators. Prior to conducting any statistical analyses, the principal investigator asked the Institutional Review Board at Texas A&M University for permission to conduct the present study following the written permission from the American Guidance Services (AGS) Research Director.

Instruments

The English version and the Spanish version of the Developmental Indicators for the Assessment of Learning – Third Edition (DIAL-3) were used in this study. The DIAL-3 consists of five screening areas: motor, concepts, language, self-help development, and social development. However, only the first three areas were used in this study. These areas require that children demonstrate fine and gross motor skills,

language skills, and awareness of cognitive concepts. Each of the three subtests yields a total scale score, which can be converted into a percentile, and consequently, into a standard score. In addition, the three subtests combine to yield a total scale score, which also can be converted into a percentile and a standard score. The psychometric properties of both of these measures were reviewed in Chapter II.

Research Questions

In order to address the research questions proposed in this study, several statistical analyses will be performed. Initially, descriptive statistics will be computed to obtain a better understanding of the data and to ensure that the data is adequate for conducting inferential statistics. The research questions are listed below for the convenience of the reader.

Research Question One

How do the six-month age-group internal consistency reliability estimates for the Spanish version of the DIAL-3 compare to the reported reliability estimates from the English standardization sample?

In order to address research question 1, estimates of internal consistency will be computed for the DIAL-3 data in each six-month age group using Cronbach's Alpha. The reliability estimates for the Spanish version of the DIAL-3 will then be compared to the reported reliability estimates using Feldt's W statistic to determine if the difference between the reliabilities is statistically significant.

Research Question Two

Do the intercorrelations among the DIAL-3 domains in the Spanish-speaking sample provide evidence to support the discriminant validity of the test?

In response to research question 2, Pearson-Product Moment correlations will be computed to assess the direction and strength of the relationship between each of the DIAL-3 domains.

Research Question Three

Do the models' results for the English standardization sample fit the standardization data for the Spanish version of the DIAL-3? If not, which models better describe the domain(s) evaluated by the test?

In response to research question 3, a confirmatory structural equation model will be created to evaluate the construct validity of the DIAL-3 for each language group specified and for the overall sample. The model will include the three factors proposed by the DIAL-3 authors: motor, language, and concepts. Each factor will have anywhere from six to seven associated test items measuring the construct depending on the age group. The path coefficients for the model will be estimated as well as several Goodness-of-Fit indices. Four Fit indices will be examined: the Goodness-of-Fit Index (GFI), the Comparative Fit Index (CFI), the Normed Fit Index (NFI), and the Root Mean Square Error of Approximation (RMSEA). The analyses will be conducted using AMOS version 4.0. Figure 1 illustrates a possible model to be tested. Exploratory factor analyses will be computed for those models which fail to yield adequate Goodness-of-Fit indices.

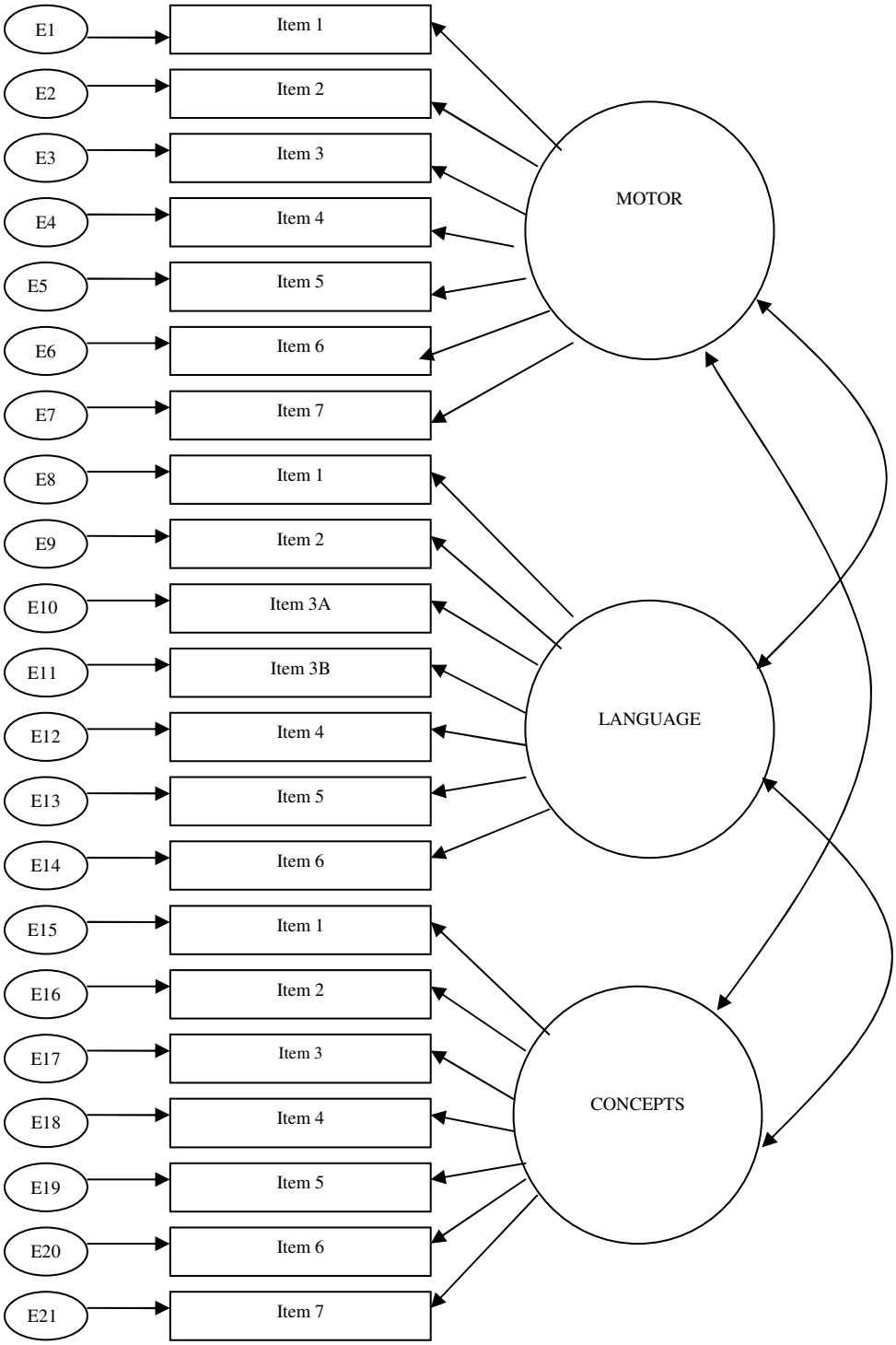

Figure 1. Confirmatory Factor Model for the DIAL-3

CHAPTER IV

STATISTICAL ANALYSES

In order to address the research questions proposed in this study, several statistical analyses were performed. Initially, descriptive statistics were computed to obtain a better understanding of the data and to ensure that the data was adequate for conducting inferential statistics. The research questions are listed below for the convenience of the reader.

Research Question One

How do the six-month age-group internal consistency reliability estimates for the Spanish version of the DIAL-3 compare to the reported reliability estimates from the English standardization sample reported in the technical manual?

In order to address research question one, estimates of internal consistency were computed for each six-month age group the Spanish standardization samples. The six-month age-groups examined were the following: 3 years 0 months to 3 years 5 months, 3 years 6 months to 3 years 11 months, 4 years 0 months to 4 years 5 months, 4 years 6 months to 4 years 11 months, 5 years 0 months to 5 years 5 months, 5 years 6 months to 5 years 11 months, 6 years 0 months to 6 years 5 months, and 6 years 6 months to 6 years 11 months. Results of the reliability analyses are provided below.

Spanish Standardization Sample

Parallel analyses were conducted in the Spanish Standardization sample. The reliability measures for the four DIAL-3 domains by six-month age group are provided in Table 1. Results indicate that the reliability estimates for the Motor domain ranged from

α = .35 to α = .77 across the age group periods examined. The Concepts domain in this sample yielded reliability estimates ranging from α = .61 to α = .90 across the age group periods examined. Internal consistency estimates of the Language domain for this sample produced estimates ranging from α = .63 to α = .90 across the age group periods examined. Lastly, Cronbach's alpha was computed for the Total domain indicating reliabilities ranging from α = .71 to α = .93 across the age group periods examined.

The aforementioned reliability estimates were compared with those obtained by the test publishers for the English version using Feldt's W statistic to determine if the differences among the reliabilities were statistically significant. Results of the performed analyses indicate that the reliability estimates were not statistically significant with the exception of the reliability comparisons in two domains of the DIAL-3 in the 3 years 0 months to 3 years 5 months age range. The difference between the reliabilities for this age bracket was found to be statistically significant for both the Motor domain ($W = 1.870$, $p < .05$) and the Language domain ($W = 2.00$, $p < .05$). A more comprehensive exploration of the comparison between the reliabilities of the Spanish version of the DIAL-3 and those reported by the test publisher for the English version will presented in Chapter V.

Table 1.

Computed Six-Month Age Group Spanish Internal Reliability Estimates

Age	N	Motor	Concepts	Language	Total
3-0 to 3-5	25	.77	.90	.90	.93
3-6 to 3-11	51	.56	.69	.63	.77
4-0 to 4-5	81	.60	.82	.78	.85
4-6 to 4-11	176	.59	.81	.70	.84
5-0 to 5-5	120	.63	.76	.69	.82
5-6 to 5-11	76	.57	.75	.69	.82
6-0 to 6-5	45	.35	.70	.64	.74
6-6 to 6-11	31	.48	.61	.72	.71

Research Question Two

Do the intercorrelations among the DIAL-3 domains in the Spanish-speaking sample provide evidence to support the discriminant validity of the test?

In response to research question 2, Pearson-Product Moment correlations were computed using the subscale standard scores to assess the direction and strength of the relationship between each of the DIAL-3 domains. The presence of low intercorrelations among the three DIAL-3 domains would be indicative of adequate discriminant validity, while high intercorrelations would be indicative of poor discriminant validity (Campbell, 1960).

Spanish Standardization Sample

The Pearson-Product Moment correlations among the three DIAL-3 domains yielded statistically significant results at $p < .01$. There was a low correlation between the Motor Domain and the Concept Domain ($r = .4$, $p < .01$) and share 16 percent of the variance. In addition, a low correlation also was found between the Motor Domain and the Language Domain ($r = .36$, $p < .01$) and share 13 percent of the variance. A correlation also was conducted between the Concepts and Language Domains, which resulted in a moderate linear relationship ($r = .61$, $p < .01$) and share 37.2 percent of the variance. The intercorrelations among the DIAL-3 domains were slightly lower for the Spanish standardization sample than for the English standardization sample. The intercorrelations among the DIAL-3 domains can be for the Spanish standardization sample can be found in Table 2.

Table 2.

Intercorrelations among DIAL-3 domains for Spanish Standardization Sample

	Motor	Concepts	Language	Total
Motor	1			
Concepts	.40 **	1		
Language	.36 **	.61 **	1	
Total	.70 **	.85 **	.82 **	1

** Correlation is significant at the 0.01 level (2-tailed).

Research Question Three

Do the models' results for the English standardization sample fit the standardization data for the Spanish version of the DIAL-3? If not, which models better describe the domain(s) evaluated by the test?

In response to research question 3, a confirmatory structural equation model was created to evaluate the construct validity of the DIAL-3 for each language group specified and for the overall sample. The model consists of the three factors proposed by the DIAL-3 authors: Motor, Language, and Concepts. Each factor tested contains only those items which are administered to every person in the standardization sample, since the statistical analyses to answer this question are contingent upon this. Items that were contingent on a given age or on the successful performance of another item within each subscale were excluded from the model. Therefore, all three factors in the different models tested only contain six associated items as opposed to the original seven items suggested by the authors. Following these criteria, item 7 from the Motor Domain, item 5 from the Language Domain, and item 3 from the Concepts Domain were excluded from each model. The path coefficients for each model were estimated as well as several Goodness-of-Fit indices. Four Fit indices also were examined: the Goodness-of-Fit Index (GFI), the Comparative Fit Index (CFI), the Normed Fit Index (NFI), and the Root Mean Square Error of Approximation (RMSEA). The analyses were conducted using AMOS version 4.0 and consist of three steps. The first step consists of constructing a baseline model that is tested separately for each group being considered. The second step consists of creating a combined model including the data from both the English and Spanish

standardization samples. The third and final step consists of creating a combined model while constraining all factor loadings, factor variances, and factor covariances. Figure 2 illustrates the theoretical model suggested by the DIAL-3 which will be tested.

According to Joreskog (1971), when testing the invariance of a particular factor structure, one must begin with a global test of the equality of covariance structures among groups. By doing this, one tests the null hypothesis that there is no difference in the population variance-covariance matrices among the groups to be tested. Significant differences among the groups provides evidence to support their nonequivalence and calls for subsequent testing of increasingly restrictive hypotheses in order to identify the source of variance that prevents the variance-covariance matrices of the different groups from being equivalent. On the other hand, acceptance of the null hypothesis would suggest that all groups are equivalent and that the data should be pooled and treated as a single group. According to Byrne (2001), when testing multigroup invariance there are three types of parameters tested. These parameters include factor loading paths, factor variances/covariances, and structural regression paths. However, in the case of confirmatory factor analyses models, the structural relations among the factors and the pattern of factor loadings are the parameters that are usually of interest (Byrne, 2001).

Before factorial invariance is tested, one must first construct a baseline model and test it separately for each group being considered (Byrne, 2001). The resulting models would represent the best fit for each group and reflects the extent to which the underlying structure fits the data without constraints (Byrne, 2001). In this study, single-group analyses were conducted for each language group (English and Spanish) using the

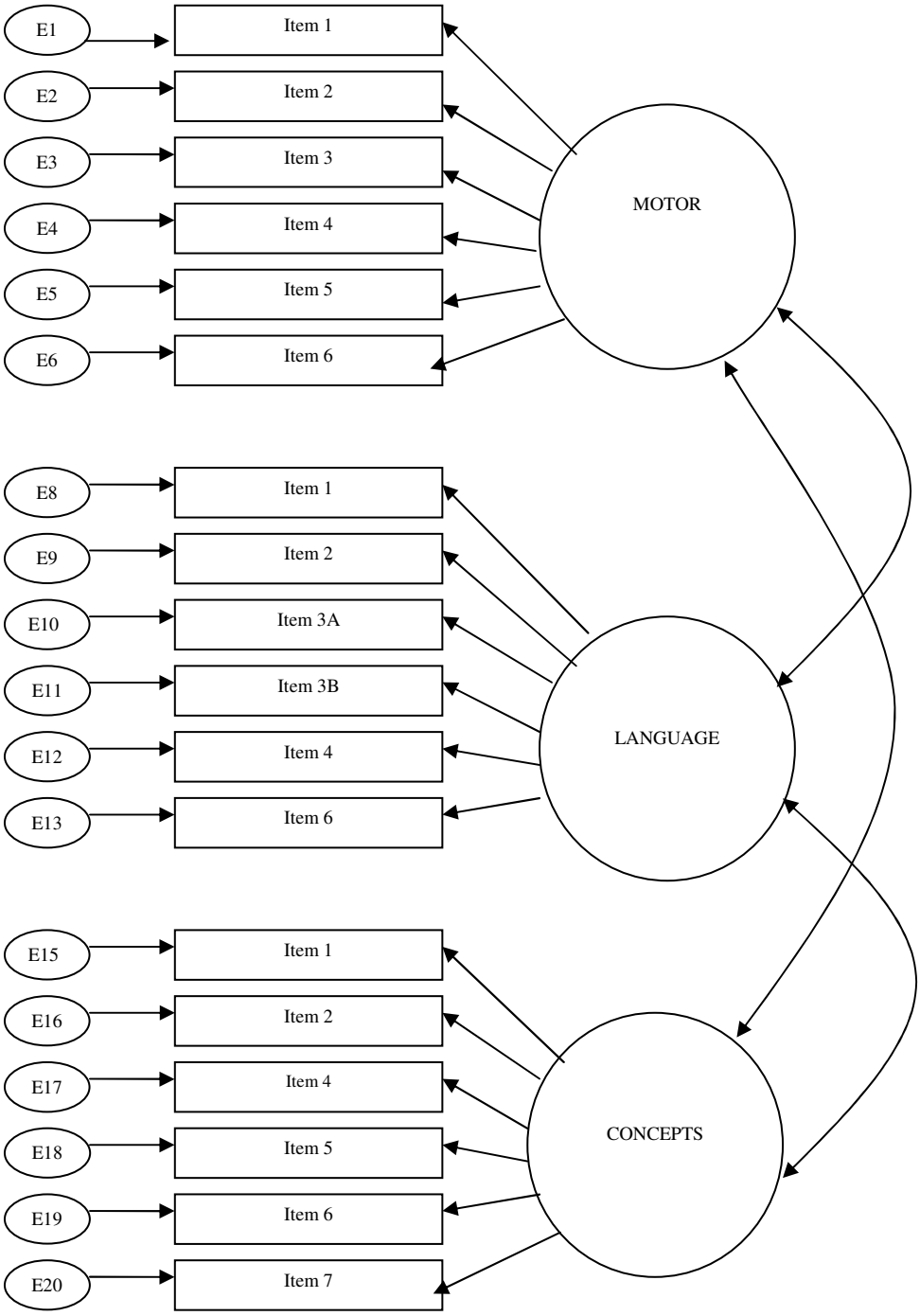

Figure 2. Proposed Model for the DIAL-3

three-factor confirmatory model suggested by the authors of the DIAL-3. Results of the English-only baseline model can be seen in Figure 3, while this same analysis for the Spanish-only model can be seen in Figure 4. The Chi-square value for the English-only model was 723.67, while the same statistic for the Spanish-only model yielded a value of 518.37. The Goodness-of-Fit indices for the English-only model indicate good model fit, suggesting that the suggested three factor model is appropriate for this analysis. Table 3 displays the Goodness-of-Fit indices for the English-only model. These same indices for the Spanish-only model were lower than those obtained for the English-only model, yet they were still acceptable, also suggesting that the proposed model is appropriate. Table 4 displays the Goodness-of-Fit indices for the Spanish only model.

A second step before testing the invariance of the parameters for the English and Spanish language groups was to create a model including the data from both groups at the same time. This second model tested the validity of the three-factor structure across the two groups simultaneously. According to Byrne (2001), the purpose of this model is to establish a baseline value against which all subsequent models are to be compared. The Chi-square value for this overall model is equal to the sum of the Chi-square value of each single-group model. Thus, the Chi-square value for this model was 1240.31. The Goodness-of-Fit indices can be found in Table 5. The structural relations among the factors and factor loading can be found in Figure 5.

The final step in this process is to construct a combined model with constraints in place (i.e. holding paths constant for the English and Spanish standardization samples). In order to test for the invariance of the parameters among the two language groups,

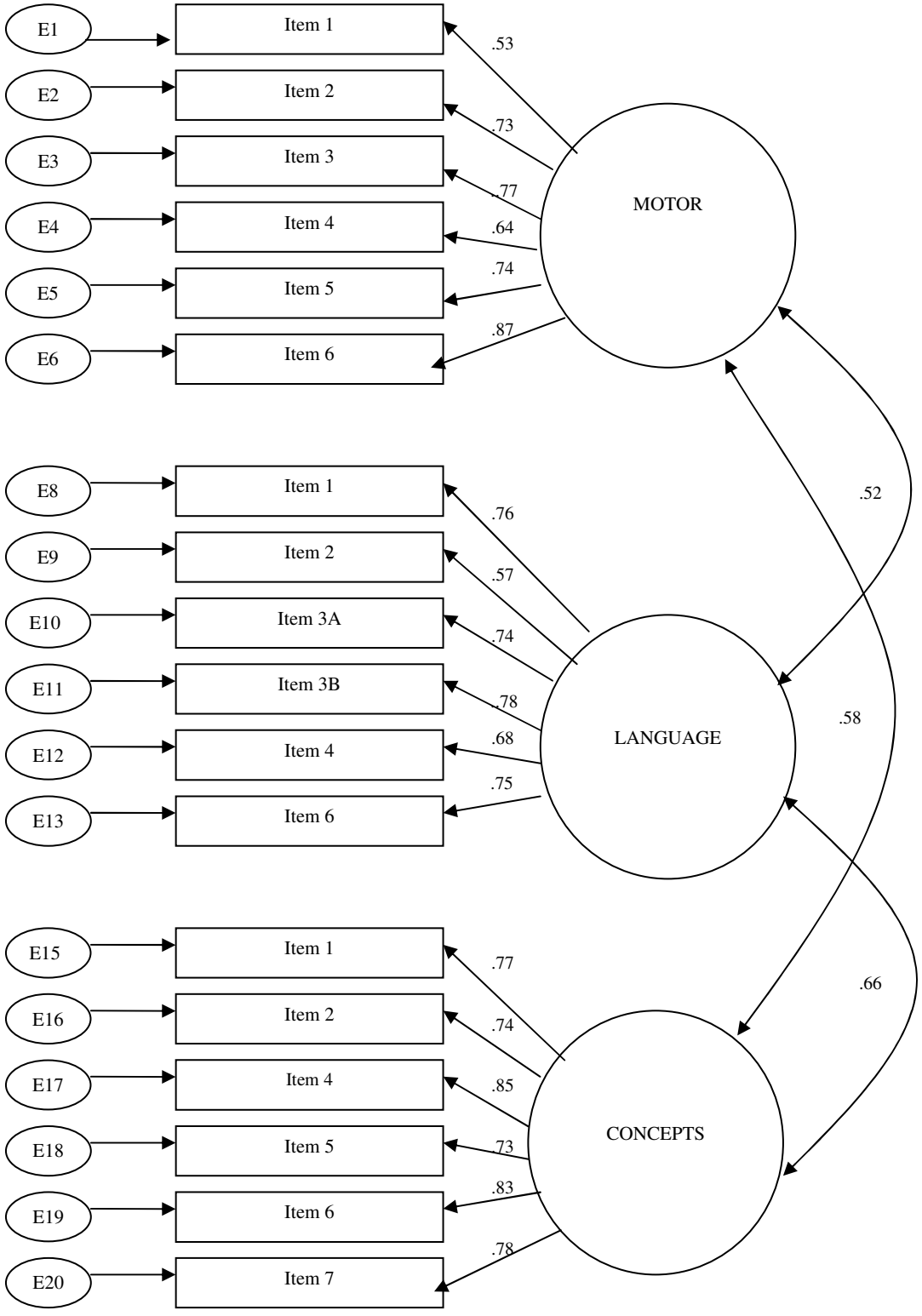

Figure 3. English Only Confirmatory Factor Analysis Model

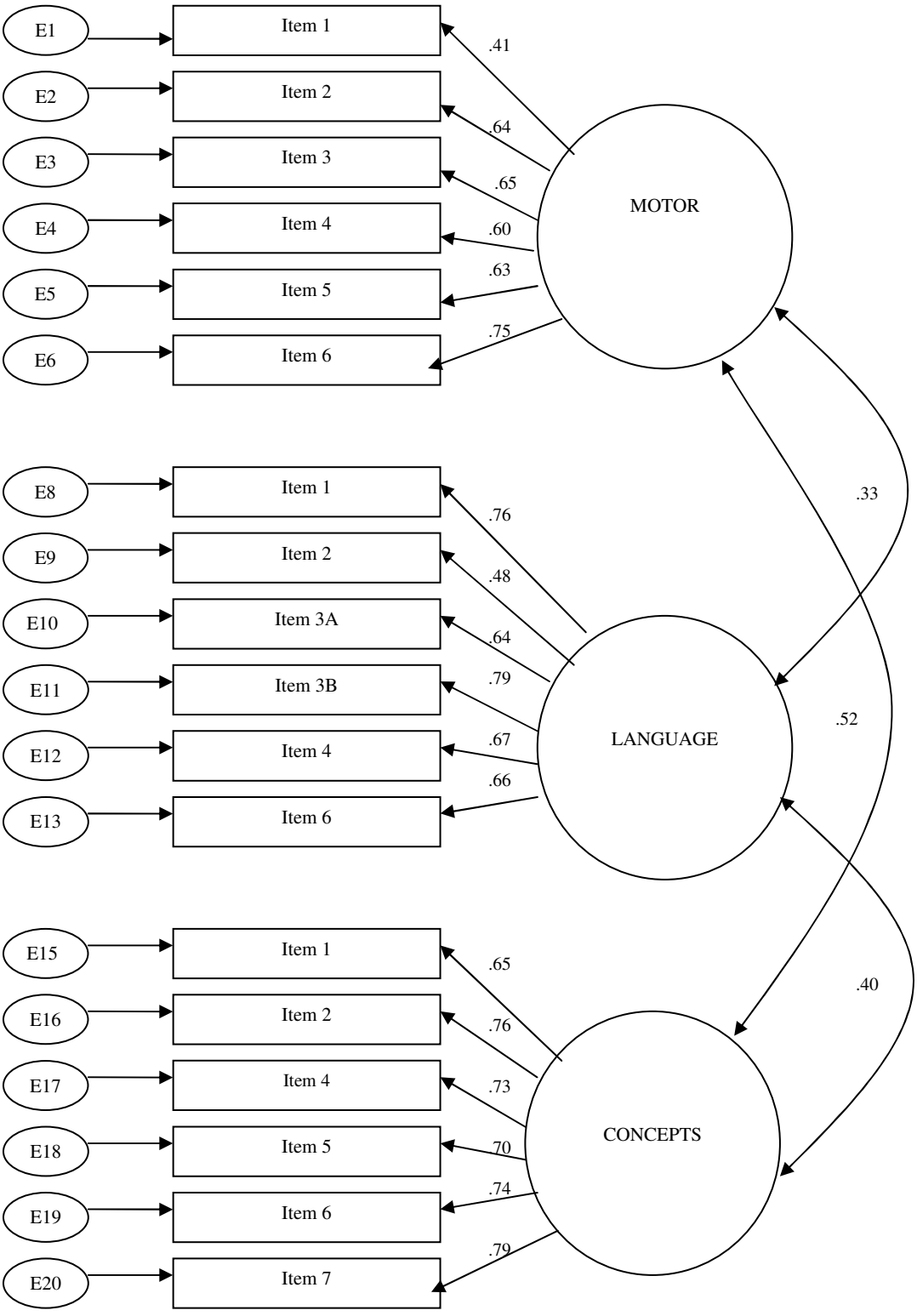

Figure 4. Spanish Only Confirmatory Factor Analysis Model

Table 3.

Goodness-of-Fit Indices for the English-only Model

	Goodness-of-Fit Index	Normed Fit Index	Comparative Fit Index	Root Mean Square Error of Approximation
Default Model	.95	.96	.97	.05
Saturated Model	1.00	1.00	1.00	
Independence Model	.19	.00	.00	.27

Table 4.

Goodness-of-Fit Indices for the Spanish-only Model

	Goodness-of-Fit Index	Normed Fit Index	Comparative Fit Index	Root Mean Square Error of Approximation
Default Model	.90	.90	.92	.07
Saturated Model	1.00	1.00	1.00	
Independence Model	.26	.00	.00	.23

Table 5.

Goodness-of-Fit Indices for the English-Spanish Model

	Goodness-of-Fit Index	Normed Fit Index	Comparative Fit Index	Root Mean Square Error of Approximation
Default Model	.93	.94	.95	.06
Saturated Model	1.00	1.00	1.00	
Independence Model	.21	.00	.00	.26

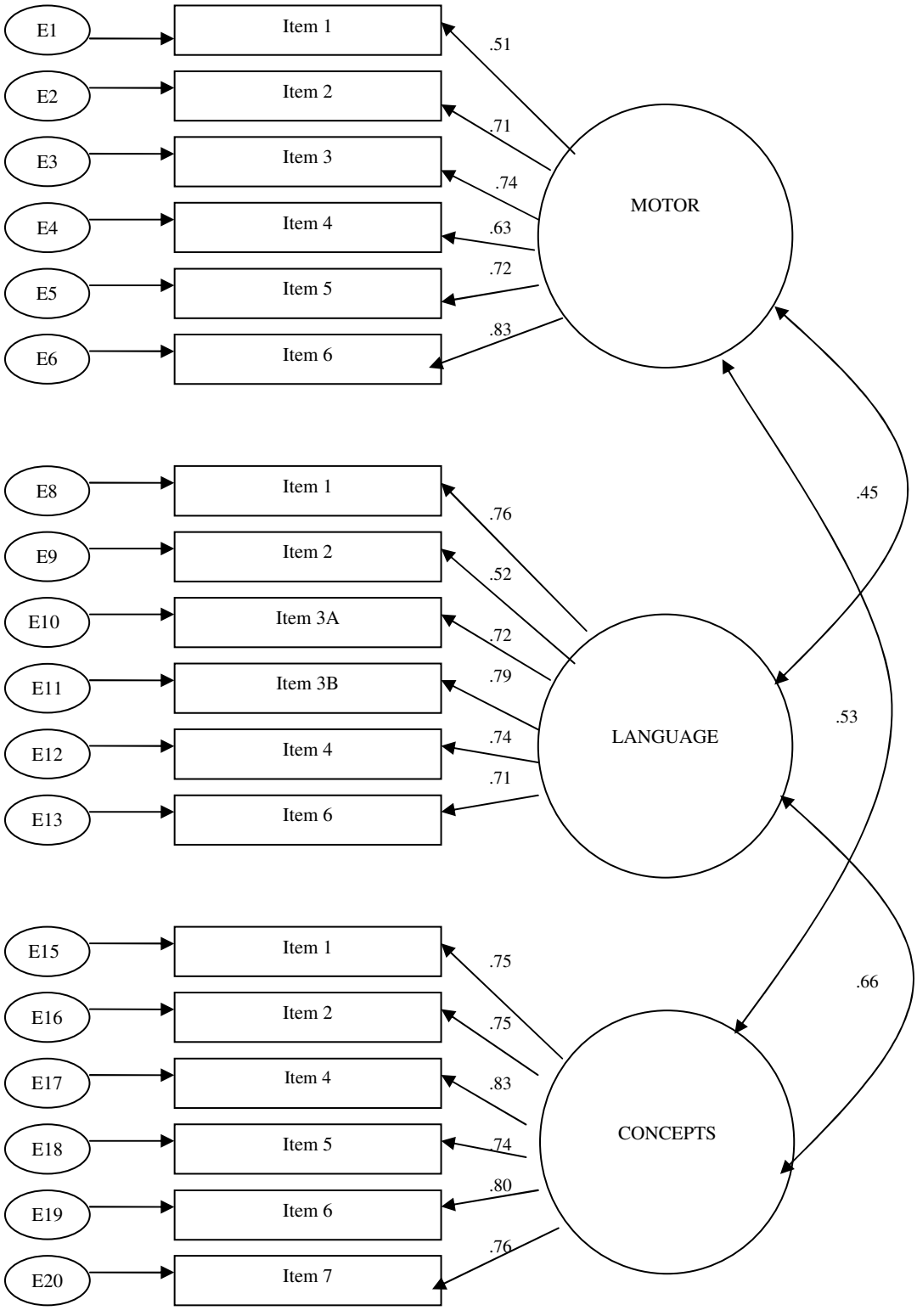

Figure 5. English-Spanish Combined Model for the DIAL-3

constraints were placed on all factor loadings, factor variance, and factor covariances. The error terms associated with each item were not constrained since doing so would represent an overly restrictive test of the data (Byrne, 2001). The Chi-square value for the constrained language model was 1404.36 with 285 degrees of freedom. The Chi-square difference between the baseline model and the constrained model yielded a value of 164.05 with 57 degrees of freedom. The critical Chi-square value at $p = .01$ with 57 degrees of freedom is 84.73. Therefore the Chi-square value for the difference between the baseline model and the constrained model is considered to be statistically significant ($p < .01$). This provides one piece of evidence to reject the null hypothesis that the variance-covariance matrices are not equal between the English and Spanish language groups. However, it has been documented that the Chi-square statistic is very sensitive to significant differences, especially with large sample sizes. Therefore, in order to fully reject this null hypothesis, one also must look the Goodness-of-Fit indices yielded by the constrained model. Table 6 shows the Goodness-of-Fit indices for the constrained model. All of the indices yielded by this model appear to show adequate Goodness-of –Fit indices providing support for the invariance among the variance-covariance matrices between the language groups. Given the adequate fit of the model, changes to the model suggested by the modification indices were not incorporated. Overall, apart from the significant Chi-square test yielded by this model (susceptible by large sample sizes), all other indices appear to indicate that the data for this test should be treated as a single group and not as separate groups. In other words, results indicate that the same

Table 6.

Goodness-of-Fit Indices for the Constrained Model

	Goodness-of-Fit Index	Normed Fit Index	Comparative Fit Index	Root Mean Square Error of Approximation
Default Model	.93	.94	.95	.04
Saturated Model	1.00	1.00	1.00	
Independence Model	.21	.00	.00	.18

underlying factor structure can be used for both the English and Spanish standardization, in this case a three-factor structure.

Summary

The psychometric properties of the DIAL-3 were explored beyond the information provided in the DIAL-3 technical manual. Particularly, the six-month age-group reliability estimates were computed for the Spanish standardization sample and were compared to the reliability estimates reported for the English standardization sample. A more detailed exploration of the reliability estimate comparisons for the English and Spanish version of the DIAL-3 will be presented in Chapter V. In addition, the discriminant validity of the instrument was evaluated by conducting correlations among the three DIAL-3 domains for both the English and Spanish standardization samples. The correlations among the three DIAL-3 domains also were comparable for both samples, although the correlations in the Spanish standardization sample were slightly lower than those from the English standardization sample. Lastly, the construct validity of the DIAL-3 also was examined through the use of a series of confirmatory factor analyses. Results of these analyses indicate that the three-factor model suggested by the DIAL-3 authors is tenable for both the English and Spanish standardization samples.

CHAPTER V

SUMMARY, DISCUSSION, AND RECOMMENDATIONS

This chapter will present a summary of the results of this study along with the inferences made based on the results. Additionally, the limitations of the study will be discussed. The implications for practice, as well as the recommendations for future research also will be provided.

Summary and Discussion

The purpose of this study was to further explore the psychometric properties of the DIAL-3 and determine its adequacy for use with both English-speaking and Spanish-speaking students, ages 3 to 6 years. The study examined the six-month age group reliability estimates for both the English and Spanish standardization samples. Moreover, the discriminant and construct validity (i.e. factorial validity) of the test also were examined.

The rationale of the study was established through a series of topics that stress the need for instruments that possess adequate psychometric properties, particularly when evaluating CLD students. The rationale for the study began with a discussion of the changing demographics in the United States and the challenges public schools encounter to address the demands of their students. In order to obtain a better understanding of the issues young children face, critical protective and risk factors for academic success were discussed. Particular emphasis was made on the fact that many children who come from culturally and linguistically diverse backgrounds do not have the privilege of relying on the same protective factors as their non-minority counterparts. The enactment of federal

laws such as IDEA, section 504, and ADA were a direct result of the need to protect the welfare of students who are at-risk, including CLD students. The Child Find process is the part of IDEA which concentrates on the early identification of students with disabilities and delineates the school's responsibility towards these students. In addition, to illustrate the need for early detection in the schools, the number of young children (ages 3 to 6 years) in special education was reported.

Screening instruments were posited as effective and efficient tools for the early detection of disabilities among students. In order to discriminate between instruments with and without adequate psychometric properties, the testing standards set forth by national organizations such as APA, NASP, and ETS were discussed. These testing standards become particularly important when attempting to accurately detect possible disabilities in culturally and linguistically diverse young children. Nevertheless, a lack of screening instruments appears to exist to address the needs of young children with CLD backgrounds. Walton and Vazquez-Nuttall (1992) found that only six screening instruments met their criteria for use with minority populations. The DIAL-3 was posited as an instrument that can be used to address the needs of both English and Spanish-speaking children

The standardization sample for both the English and Spanish versions of the test were obtained through AGS. Explicit permission to use this data was obtained from both the publishing company as well as the Institutional Review Board at Texas A&M University. The standardization sample included 1,560 children for the English version of the DIAL-3. The standardization sample for the Spanish version of the test consisted of

536 monolingual Spanish-speaking children in the U.S., 26 children from Panama, and 43 children from Puerto Rico for a total sample size of 605 children. All statistical analyses for this study were conducted using standard scores.

The six-month age-group internal consistency reliability estimates for the Spanish version of the DIAL-3 were computed in order to compare them to the reliability estimates of the English standardization sample reported in the DIAL-3 technical manual. According to Murphy and Davidshofer (2001) reliability estimates ranging from .70 to .79 indicate low reliability, estimates ranging from .80 to .89 indicate moderate reliability, and estimates ranging from .90 to .99 indicate high reliability. Alfonso and Flanagan (1999) and Lehr, Ysseldyke, and Thurlow (1987) provide evidence to show that an instrument that is designed for screening is adequate if its reliabilities are .80 or higher. The six-month age-group reliability estimates computed for the Spanish version of the DIAL-3 were similar to those of the English version of the test. For the Motor Domain in the Spanish version of the instrument, the internal reliability estimates ranged from .35 to .77, while for the English version these same reliability estimates ranged from .45 to .74. Overall, the differences between the reliability estimates of the Motor Domain for the Spanish and English versions were not statistically significant for any age range with the exception of the 3 years 0 months to 3 years 5 months, which was found to be statistically significant ($W = 1.87$, $p < .05$). The alpha coefficients for the Motor Domain are presented in Table 7 for both versions of the DIAL-3. For the Concepts Domain in the Spanish version of the instrument, the internal reliability estimates ranged from .61 to .90, while for the English version these same reliability estimates ranged from .65 to .88.

Table 7.

Internal Reliability Estimate Comparison for Motor Domain by Six-Month Age Group

	Spanish Version		English Version		
Age	N	Motor	N	Motor	W
3-0 to 3-5	25	.77	110	.57	1.870 **
3-6 to 3-11	51	.56	197	.54	1.045
4-0 to 4-5	81	.60	233	.63	0.925
4-6 to 4-11	176	.59	297	.68	0.780
5-0 to 5-5	120	.63	273	.71	0.784
5-6 to 5-11	76	.57	165	.74	0.605
6-0 to 6-5	45	.35	162	.69	0.477
6-6 to 6-11	31	.48	121	.45	1.058

** Reliability difference is significant at the 0.05 level

The differences between the reliability estimates of the Concepts Domain for the Spanish and English versions were not statistically significant for any age range. The alpha coefficients for the Concepts Domain are presented in Table 8 for both versions of the instrument. For the Language Domain in the Spanish version of the instrument, the internal reliability estimates ranged from .63 to .90, while for the English version these same reliability estimates ranged from .66 to .80. As in the Motor Domain, the difference between the reliability estimates of the Language Domain for the Spanish and English versions were not statistically significant for any age range with the exception of the 3 years 0 months to 3 years 5 months, which was found to be statistically significant ($W = 2.00$, $p < .05$).The alpha coefficients for the Language Domain are presented in Table 9 for both versions of the DIAL-3. For the Total Domain in the Spanish version of the instrument, the internal reliability estimates ranged from .71 to .93, while for the English version these same reliability estimates ranged from .77 to .90. No statistically significant differences were found between the reliabilities of the English and Spanish versions for the Total Domain of the DIAL-3. The alpha coefficients for the Total Domain are presented in Table 10.

In general, the reliability estimates for both versions of the test mostly range from low to moderate reliability. The reliability estimates for the Motor Domain in both the English and Spanish versions of the test appear to be the weakest for all age groups. Statistically significant differences were found between the reliability estimates of the English and Spanish versions for both the Motor and Language Domains in the 3 years 0

Table 8.

Internal Reliability Estimate Comparison for Concepts Domain by Six-Month Age Group

	Spanish Version		English Version		
Age	N	Concepts	N	Concepts	W
3-0 to 3-5	25	.90	110	.88	1.200
3-6 to 3-11	51	.69	197	.86	0.452
4-0 to 4-5	81	.82	233	.85	0.833
4-6 to 4-11	176	.81	297	.81	1.000
5-0 to 5-5	120	.76	273	.85	0.625
5-6 to 5-11	76	.75	165	.82	0.720
6-0 to 6-5	45	.70	162	.65	1.167
6-6 to 6-11	31	.61	121	.67	0.846

Table 9.

Internal Reliability Estimate Comparison for Language Domain by Six-Month Age Group

	Spanish Version		English Version		
Age	N	Language	N	Language	W
3-0 to 3-5	25	.90	110	.80	2.000 **
3-6 to 3-11	51	.63	197	.76	0.649
4-0 to 4-5	81	.78	233	.74	1.182
4-6 to 4-11	176	.70	297	.66	1.133
5-0 to 5-5	120	.69	273	.79	0.677
5-6 to 5-11	76	.69	165	.80	0.645
6-0 to 6-5	45	.64	162	.77	0.639
6-6 to 6-11	31	.72	121	.77	0.821

** Reliability difference is significant at the 0.05 level

Table 10.

Internal Reliability Estimate Comparison for Total Domain by Six-Month Age Group

	Spanish Version		English Version		
Age	N	Total	N	Total	W
3-0 to 3-5	25	.93	110	.89	1.571
3-6 to 3-11	51	.77	197	.87	0.565
4-0 to 4-5	81	.85	233	.86	0.933
4-6 to 4-11	176	.84	297	.85	0.938
5-0 to 5-5	120	.82	273	.90	0.556
5-6 to 5-11	76	.82	165	.89	0.611
6-0 to 6-5	45	.74	162	.85	0.577
6-6 to 6-11	31	.71	121	.77	0.793

months to 3 years 5 months age range. This appears to indicate that the items within these two domains do not consistently measure the purported area in an equal manner for this particular age range. Results indicate that the reliability estimates for the Spanish version of the Motor and Language domains were significantly higher. The direction of the difference between the reliabilities in the 3 years 0 months to 3 years 5 months age range in both the Motor and Language domains indicates that the scores for the Spanish-speaking sample were actually more reliable than those in the English-speaking sample.

Overall, given the nature and age range that the instrument is designed to assess, the obtained reliability estimates fall within the expected range. Because young children develop abilities at different paces, it is difficult to measure any particular area with a high degree of reliability. For example, in one month time children may make great gains in physical development and may not meet any new academic milestones and in the following month, this pattern may reverse. Nevertheless, most commonly used screening, readiness, and behavioral instruments for this age group have been criticized for having questionable validity and/or reliability (Carney & Merrell, 2002; Rafoth, 1997; Crnic & Lamberty, 1994; Harrison, 1991; Bredekamp & Shepard, 1989; Bracken, 1987; Meisels, 1987). Therefore, given this information, it appears that the reliability estimates obtained for the DIAL-3 are common for this age range, but as with all screening instruments, need to be interpreted with caution.

Intercorrelations among the DIAL-3 domains were conducted for the Spanish version of the test to examine the discriminant validity of the test and then compared them to those obtained for the English version by the test publishers. The correlations

among the different domains were generally low for both versions of the test, with the exception of the correlation between the Concepts and Language domains. The correlation between the Motor Domain and the Concepts Domain was .40 for the Spanish version of the test and .50 for the English version. The correlation between the Motor Domain and the Language Domain for the Spanish version of the DIAL-3 was .36, while this same correlation for the English version was .41. The correlation between the Concepts Domain and the Language Domain was .61 for the Spanish version of the instrument and .65 for the English version.

The correlation between the Concepts Domain and the Language Domain was moderate in both versions of the test, and suggests a reduced differentiation of the Concepts and Language constructs. This reduced ability to differentiate between these two constructs may be partly due to interwoven nature of the domains. By definition, a concept is something that is conceived in the mind. In other words, an abstract or general idea inferred or derived from specific instances. In order for an examiner to request a task of the examinee, the examiner typically uses language to express his/her request. Furthermore, in order for the child to demonstrate knowledge of a concept, he/she usually does this through the use of language. So, in essence, as the examiner evaluates a child's ability to recognize or recall concepts, the examiner also is indirectly assessing the child's ability to produce language. The same idea holds when assessing language in a child. In assessing the language of the child, he/she also must recognize and understand the concepts being asked by the examiner. For instance, a child may be asked for the description or function of an object in order to assess the accuracy and elaboration of the

response. In order to do this, not only must the child first know the concept that he/she is being asked to describe, but he/she also must know the concepts imbedded in the description of the object. To illustrate this notion, consider asking a child what the function of a stove is. A child who knew what the concept of a stove was would most likely answer that it is something that is used to heat food. In analyzing the response to this question, the child not only had to know what a stove was but also had to know the concepts of heat and food. Therefore, it appears that both the Concepts and Language Domains are interconnected by definition, which explains the degree of correlation found between the two. Given the nature of these constructs, the resulting degree of intercorrelation found between the Concepts and Language Domain seems to be acceptable. All in all, the results obtained in this study appear to support the discriminant validity of the test.

In addition to the six-month age-group reliability estimates and an examination of the discriminant validity for both the English and Spanish versions of the test, the construct validity of the test also was explored. A series of confirmatory factor analyses were conducted in order to determine the invariance of the variance-covariance matrices between the English and Spanish standardization samples. The Goodness-of-Fit Index for the constrained model was .93. The Normed Fit Index for the same model yielded a value of .94. The Comparative Fit Index for this model was .95. Lastly, the Root Mean Square Error of Approximation for this model yielded a value of .04. These results indicate that the three-factor model originally proposed by Mardell-Czudnowski and Goldenberg (1998) was adequate for both the English and Spanish versions of the DIAL-3.

Given the information provided in the technical manual in conjunction with the results obtained in this study, it appears that the DIAL-3 possesses adequate psychometric properties for screening young children with either English or Spanish-speaking backgrounds.

Limitations of the Study

This study has several limitations. One limitation is that the accessible population from which the sample was obtained may not generalize to the target population. Approximately 11.4% of the children who were tested for the Spanish language group came from a U.S. territory or another Spanish-speaking country. For this reason, the results obtained from this study may not generalize to the entire Spanish-speaking population in the United States as was the intent of the authors. A second limitation was the smaller sample size of the children in the Spanish-speaking language group when compared to the English-speaking language group. The Spanish language group was approximately one-third the size of the English language group. The inferences drawn from this study may be suspect since lower sample sizes tend to yield greater margins of sampling error. A third limitation to the study was that no information was available regarding the language proficiency of the Spanish-speaking standardization sample. For example, there was no information on whether the children included in the Spanish-speaking standardization sample were monolingual Spanish-speakers or bilingual English/Spanish-speakers. This information also hinders the generalization of the inferences drawn from this study to the Spanish-speaking population at-large.

Recommendations for Future Research

Additional research in this area should focus on the development and standardization of instruments that can accurately assess the needs of young children from culturally and linguistically diverse backgrounds. Given the growing need to assess these children, it is imperative that the utilized instruments possess the adequate psychometric properties. Furthermore, many screening instruments that have been developed in the field are currently standardized on a lower than optimal number of children from CLD backgrounds, and yet, these instruments are often used to assess minority students with the same degree of confidence as with their non-minority counterparts. Extensive psychometric analyses need to be conducted on these instruments to ensure that they are appropriate for the population which they are attempting to assess.

With this in mind, it will be important for future research to be conducted on the DIAL-3. The criterion-related validity of the DIAL-3 should be further examined by administering the instrument to young children and correlating the obtained scores with some measure of their performance in the schools. Conclusions from the results of this study indicated that the proposed factor model was appropriate for English and Spanish standardization samples. However, given the lower number of Spanish-speakers in the standardization sample, it will be interesting to determine if the factor structure continues to hold when approximately equal sample sizes are used. Moreover, studies that analyze the factor structure for each individual age group (e.g., 3-year-olds, 4 year-olds, 5 year-olds, and 6 year-olds) should be examined. In this study, the Concepts and Language Domains appeared to have had a moderate correlation. It is possible that the correlation

between the constructs may be higher for the younger children since these constructs are more difficult to differentiate. In such a case, a two-factor model that unifies the Concepts and Language Domains may be more appropriate.

Implications for Practice

Results of this study coupled with the psychometric properties provided in the DIAL-3 technical manual indicate that the DIAL-3 is an appropriate screening instrument to use when assessing the needs of young children from culturally and linguistically diverse backgrounds. The DIAL-3 assesses children in five areas of early childhood development: Motor, Concepts, Language, Self-Help Development, and Social Development. Given the brevity of the instrument, it is a useful measure for school psychologists to use in trying to address the demands of the children in the public school system. Also, the DIAL-3 is a cost effective alternative to other assessment methods which require more resources to conduct. Since this instrument is easy to use, score, and interpret, schools should be encouraged to use it as an integral part of their screening programs.

REFERENCES

Alfonso, V. C. & Flanagan, D. P. (1999). Assessment of cognitive functioning in preschoolers. In E. V. Nutall, I. Romero, & J. Kalesnik (Eds.), *Assessing and screening preschoolers* (2nd ed., pp. 186-217). New York, NY: Allyn & Bacon.

American Psychological Association (APA). (1993). Guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations. *American Psychologist, 48* (1), 45-48.

Americans with Disabilities Act of 1990, Pub. L. No. 101-336, §2, 104 Stat. 328 (1991).

Bettes, B. A. (1988). Maternal depression and motherese: Temporal and intonational features. *Child Development, 59* (4), 1089-1096.

Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment*, *5*, 313-326.

Bredekamp, S. & Shepard, L. (1989). How best to protect children from inappropriate school expectations, practices, and policies. *Young Children, 44* (3), 14–24.

Byrne, B. M. (2001). *Structural equation modeling with AMOS, basic concepts, applications, and programming.* Hillsdale, NJ: Erlbaum.

Campbell, D.T. (1960). Recommendations for the APA test standards regarding construct, trait, and discriminant validity. *American Psychologist, 15*, 546-553.

Campbell, F. A., & Ramey, C. T. (1995). Cognitive and school outcomes for high risk African American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal, 32*(4), 743-772.

Carney, A. G., & Merrell, K. W. (2002). Reliability and comparability of a Spanish-language form of the Preschool and Kindergarten Behavior Scales. *Psychology in the Schools, 39,* 367-373.

Chapa, J., & Valencia, R. (1993). Latino population growth, demographic characteristics, and educational stagnation: An examination of recent trends. *Hispanic Journal of Behavioral Sciences, 15*, 165-187.

Chen, T**,** Wang, J**.,** Mardell-Czudnowski, C.**,** Goldenberg, D., & Elliott, C. (2000). The development of the Spanish version of the Developmental Indicators for the Assessment of Learning-third edition (DIAL-3). *Journal of Psychoeducational Assessment*, *18* (4), 316-343.

Crnic, K., & Lamberty G. (1994). Reconsidering school readiness: Conceptual and applied perspectives. *Early Education and Development, 5* (2), 99-105.

Davidson, R.G. (1993). Oral preparation for literacy: Mothers' and fathers' conversations with precocious readers. (Doctoral dissertation, Graduate School of Education of Harvard University, 1993) *Dissertation Abstracts International*, *54* (5-A), 1665.

Edin, K., & Lein, L. (1997). *Making ends meet: How single mothers survive welfare and low-wage work*. New York: Russell Sage Foundation.

Educational Testing Service (ETS). (2000). *ETS standards for quality and fairness.* Princeton, NJ: Educational Testing Service.

Guilford, J. P. (1954). *Psychometric methods.* New York: McGraw-Hill

Harrison, P.L. (1991). Assessment of adaptive behavior. In B.A. Bracken (Ed.), *The*

psychoeducational assessment of young children (pp. 317–340). Needham, MA: Allyn & Bacon.

Hart, B., & Risley, T. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology, 28*, 1096-1105.

Hart, B., & Risley, T. (1995). *Meaningful differences in everyday parenting and intellectual development in young American children*. Baltimore, MD: Brookes.

Hoff, E., Laursen, B., & Tardiff, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting* (2nd ed., pp. 231–252). Mahwah, NJ: Erlbaum.

Individuals with Disabilities Education Act Amendments of 1997, Pub. L. No. 105-17, (1997).

Joint Committee on Testing Practices. (2003). *Code of fair testing practices in education* [Online]. Available at: http://www.apa.org/science/FinalCode.pdf

Joreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*, 409-426.

Kindler, A.L. (2002). *Survey of the states' limited English proficient students and available educational programs and services: 2000–2001 summary report*. Washington, DC: U.S. Department of Education, Office of English Language Acquisition, Language Enhancement and Academic Achievement for Limited English Proficient Students.

Lazar, I., Darlington, R., Murray, H., Royce, J., & Snipper, A. (1982). Lasting effects of

early education: A report from the consortium for longitudinal studies. *Monographs of the Society for Research in Child Development, 47* (Serial No. 195).

Lehr, C. A., Ysseldyke, J. E., & Thurlow, M. L. (1987). Assessment practices in model early childhood special education programs. *Psychology in the Schools, 24*, 390-399.

Lidz, C. S. (2003). *Early childhood assessment.* Hoboken, NJ: John Wiley & Sons.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. *Developmental Psychology, 36,* 596-613.

Mardell-Czudnowski, C. & Goldenberg, D. S. (1998). *DIAL-3: Developmental Indicators for the Assessment of Learning-Third Edition.* Circle Pines, MN: American Guidance Service.

Marsella, A. J., & Kameoka, V. A. (1989). *Ethnocultural issues in the assessment of psychopathology*. Washington DC: American Psychiatric Press.

McLanahan, S., & Sandefur, G. (1994). *Growing up with a single parent: What hurts, what helps*. Cambridge, MA: Harvard University Press.

Meisels, S.J. (1985). *Developmental screening in early childhood: A guide* (Rev. ed). Washington, DC: National Association for the Education of Young Children.

Meisels, S.J. (1987). Uses and abuses of developmental screening and school readiness testing. *Young Children, 42* (4-6), 68-73.

Murphy, K. R., & Davidshofer, C. O. (2001). *Psychological testing: Principles and*

applications (5th ed.). Upper Saddle River, NJ: Prentice-Hall

National Research Council (NRC). (1998). Preventing reading difficulties in young children. Committee on the Prevention of Reading Difficulties in Young Children, Commission on Behavioral and Social Science and Education, C.E. Snow, M.S. Burns, and P. Griffin (Eds.). Washington, DC: National Academy Press.

National Research Council (NRC). (2000). Family Resources. From neurons to neighboorhoods. Committee on Integrating the Sciences of Early Childhood Development, Board on Children, Youth and Families, J.P. Shonkoff, and D.A. Phillips (Eds.). Washington, DC: National Academy Press.

Portes, A., & MacLeod, D. (1996). What shall I call myself? Hispanic identity formation in the second generation. *Ethnic and Racial Studies, 19* (3), 523-547.

Rafoth, M.A. (1997). Guidelines for developing screening programs. *Psychology in the Schools, 34* (2), 129–137.

Ramey, S. L., & Ramey, C. T. (1999). The transition to school for "at-risk" children. In R. C. Pianta & M. J. Cox, (Eds.), *The transition to kindergarten,* (pp. 217-251). Baltimore, MD: Brookes.

Rebell-Britto, P., & Brooks-Gunn, J. (Eds.). (2001). *The role of family literacy environments in promoting young children's emerging literacy skills.* San Francisco, CA: Jossey-Bass.

Rehabilitation Act of 1973, 29 U.S.C. § 794; Regulations at 34 C.F.R. § 104.

Reid, R., Maag, J. W., & Vasa, S. F. (1994). Attention deficit hyperactivity disorder as a disability category: A critique. *Exceptional Children, 60*, 198-214.

Reid, R. (1995). Assessment of ADHD with culturally different groups: The use of behavioral rating scales. *School Psychology Review, 24* (4), 537-561.

Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2001, June 1). *Age 21 cost-benefit analysis of the Title I Chicago child-parent center program: Executive summary.* Retrieved November 6, 2005 from http://www.waisman.wisc.edu/cls/cbaexecsum4.html

Rogers, M., Ingraham, C., Bursztyn, A. M., Cajigas-Segredo, N., Esquivel, G., Hess, R., et al. (1999). Providing psychological services to racially, ethnically, culturally, and linguistically diverse individuals in the schools. *School Psychology International, 20* (3), 243-264.

Romero, I. (1999). Individual assessment procedures with preschool children. In E. Nuttal, I. Romero, & J. Kalesnik (Eds.), *Assessing and screening preschoolers* (pp. 59-71). Needham Heights, MA: Allyn & Bacon.

Ross, D. M. & Ross, S. A. (1982). *Hyperactivity: Current issues, research and theory.* (2nd ed.). New York, NY: Wiley & Sons

Salisbury, C.L. & Smith, B. (1993). Effective practices for preparing young children with disabilities for school. (Report No. E513) ERIC Document Reproduction Service No. ED358 675.

Schweinhart, L. J., & Weikart, D. P. (1997). The High/Scope Preschool Curriculum comparison study through age 23. *Early Childhood Research Quarterly, 12* (2), 117-143.

Smith, J. R., Brooks-Gunn, J., & Klebanov, P. K. (1997). The consequences of living in

poverty for young children's cognitive and verbal ability and early school achievement. In G. J. Duncan & J. Brooks-Gunn (Eds.), *Consequences of growing up poor*. (pp. 132-189). New York, NY : Russell Sage Foundation Press.

Snow, C. E., & Paez, M. M. (2004). The Head Start classroom as an oral language environment: What should the performance standards be? In E. Zigler & S. J. Styfco (Eds.), *The Head Start debates* (pp. 113-128). Baltimore, MD: Brookes.

Storch, S. A. & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal model. *Developmental Psychology, 38*, 934-947.

Thurlow, M. L. (1992). Issues in the screening of preschool children. In E.Vazquez-Nuttal, I. Romero, & J. Kalesnik (Eds.), *Assessing and screening preschoolers: Psychological and educational dimensions,* 2nd ed. (pp. 67-82). Boston, MA: Allyn & Bacon.

U.S. Bureau of the Census. (1994). Household and family characteristics, (*Current Population Reports, P20-484*). Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census. (1999). School enrollment in the United States – social and economic characteristics of students. *(Current Population Reports, P20-533).* Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census. (2002). *Current Population Survey.* March Supplement. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. (2003). *24th annual report to Congress on the*

implementation of the IDEA. Washington DC: US Government Printing Office.

U.S. General Accounting Office (1998). Head Start programs: Participants characteristics, services, and funding. (*Report No. HEHS 97-59*) Washington, DC: Author.

Walton, J. R., & Vazquez-Nuttall, E. (1992). Preschool evaluation of culturally different children. In E. Vazquez-Nuttall, I. Romero, & J. Kalesnik (Eds.). (1992). *Assessing and screening preschoolers: Psychological and educational dimensions* (pp. 281-300). Needham, MA: Allyn & Bacon.

Whitehurst, G. J., & Fischel, J. E. (2000). A developmental model of reading and language impairments arising in conditions of economic poverty. In D. Bishop & L. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 53-71). East Sussex: Psychology Press.

Winquist-Nord, C., & Griffin, J. A. (1999). Educational profile of 3- to 8-year-old children of immigrants. In D. Hernandez (Ed.), *Children of immigrants: Health, adjustment, and public assistance*. (pp. 348-409). Washington, DC: National Academy Press.

Zill, N., Collins, M., West, J. & Hausken, E. (1995). *Approaching Kindergarten: A Look at Preschoolers in the United States*, (NCES publication No. 95-280). Washington, DC: U.S. Department of Education.

Zill, N. (1996). Family change and student achievement: What have we learned, what it

means for schools. In A. Booth and J. F. Dunn (Eds.), *Family school links: How do they effect educational outcomes?* Hillsdale, NJ: Lawrence Erlbaum Associates.

VITA

Name:	Roman Garcia de Alba
Address:	University of Texas Pan American, Department of Educational Psychology, 1201 W. University Drive, Edinburg, Texas 78541
Email Address:	spsy2005@yahoo.com
Education:	B.A., Psychology, Texas A&M University, 2000 Ph.D., School Psychology, Texas A&M University, 2006