

Reading Results: A Critical Look at Standardized Testing and the Linguistic Minority

Joanie Shannon

Submitted in Partial Fulfillment of the Requirements for the Degree

Master of Science in Education

School of Education  
Dominican University of California

San Rafael, CA

August 2008

### ACKNOWLEDGEMENTS

I would like to acknowledge my professors at Dominican University of California for the wisdom they have offered me and the unconditional support they have given. I would also like to thank my colleagues for their inspirational ideas and amazing insight on topics that seemed abstract at times. I would like to acknowledge the amazing support from my family both financially and emotionally. You have been a constant support system in my life and I cannot thank you enough.

TABLE OF CONTENTS

TITLE PAGE ..... 1

ACKNOWLEDGEMENTS..... 2

TABLE OF CONTENTS ..... 3

ABSTRACT ..... 4

INTRODUCTION..... 5

    PROBLEM..... 6

    PURPOSE STATEMENT..... 6

    RESEARCH QUESTION..... 6

THEORETICAL RATIONALE..... 8

    ASSUMPTIONS ..... 9

    BACKGROUND AND NEED..... 9

REVIEW OF THE LITERATURE ..... 11

    SOCIOECONOMIC STATUS AND THE ATTAINMENT OF EDUCATION ..... 12

    STANDARDIZED TESTING AND THE CORRELATION TO LANGUAGE ..... 18

    STANDARDIZED TESTING, NO CHILD LEFT BEHIND AND MINORITIES ..... 24

    SUMMARY OF MAJOR FINDINGS ..... 35

    LIMITATIONS AND GAPS IN THE RESEARCH ..... 36

    IMPLICATIONS FOR FUTURE RESEARCH ..... 37

    OVERALL SIGNIFICANCE OF THE LITERATURE..... 38

PILOT STUDY ..... 40

    METHOD..... 40

*Sample and Site*..... 40

*Access and Permissions* ..... 40

*Data Gathering Strategies* ..... 41

*Data Analysis Approach* ..... 41

*Ethical Standards*..... 41

*Results or Findings* ..... 42

*Analysis of Themes*..... 43

    SUMMARY OF MAJOR FINDINGS ..... 44

*Comparison of Findings to Existing Literature* ..... 44

*Limitations of the Study* ..... 45

*Implications for Future Research* ..... 45

*Overall Significance of the Study*..... 46

CONCLUSIONS ..... 48

REFERENCES..... 51

### ABSTRACT

A critical look into assessing the Standardized Test and Reporting (STAR) test data among English language learners gives educators a chance to examine the cultural biases present within the standardized test movement started by the No Child Left Behind Act of 2001. In particular, The STAR test results seem to reflect that the test is geared toward students who not only speak the mainstream language but who grew up in the mainstream culture.

After examining past student's STAR results, there seemed to be a discrepancy between the test score and the student's ability. This also seemed to be integrated with the student's cultural background, when they came to the United States and how long they have lived in the country. A student's level of language acquisition is considered when their test results are gauged but there seems to be no consideration toward whether or not they are acculturated.

After reviewing literature on the topic, it seems as though a plan for intervention must be set in place if any changes are to be made. The literature shows that most standardized tests do have an underlying bias, and also reveals which cultures are the most susceptible to biases within tests.

The pilot study conducted in line with the research shows that students from the Latino culture in particular fall victim to test bias. A possible plan for intervention may be a small, localized step toward developing a plan for success for language learners.

## INTRODUCTION

After an in-depth look at the data and test scores from the Standardized Test and Reporting (STAR) given to middle school students in a moderate income district, a gap in test scores became apparent. This gap occurred among the test scores of the English language learners.

Some of them, regardless of their language acquisition level, would score “far below basic” on the STAR and some would score “below basic” or “basic.” Scores were also dropping among students who were living in California and who were becoming mainstreamed in school. Therefore, this did not seem to be determined by language skills alone, which left me with many questions regarding how the results were read and what was considered in the reading.

I was aware that these students were put into a category of English language learners (ELLs) and that this was considered when their results were read, but I realized there was something more to the test scores than just the level of English acquisition. Students that I had taught in class, with average language skills were scoring extremely low on a test that measures basic skills. Also, students who were placed in mainstream classes with primary English speakers, were scoring lower each year although they were progressing in their language acquisition.

This problem seemed to be apparent district-wide. As of now, there is neither a solution nor a plan for intervention. The plan that had been in place was deemed unsuccessful and dropped altogether. I believe that the gap can be bridged with a critical look at why these students are doing poorly on this particular test, why their scores are going down as their skills are coming up and how this can be rescinded.

### Problem

The problem is determining the criteria that measures why language learners are scoring low on standardized tests. If it became apparent that the tests had embedded cultural biases, then the test could be revamped to portray an accurate measure of subject-matter mastery among students whose primary language is not English.

ELLs are put into a category of “not English proficient” without any further investigation into why their scores are low. Outside elements, besides language, may affect how they score on their STAR. These low scores could have to do with socioeconomic status and dialect in their home language and in their home culture. A solution to this problem would be to find a way to assess the results intersectionally, a variety of ways that tie together, to find out what is really bringing their scores down. Once this information has been determined, a plan to help these students become successful could be put into place.

### Purpose Statement

The purpose of this study is to review the research that examines how school officials can assess test results in a varied way that determines the multiple reasons ELLs are scoring “below basic” and “far below basic” on the STAR. This review of the research may give *educationists* (Spring, 2005) an insight into why certain language learners are scoring below basic on their STAR while others are not. Once they know why this is occurring, a plan for intervention and eventually a solution to the problem can be formed.

### Research Question

This brings us to the important question that must be addressed as part of the literature review process. The question directs the research model and narrows the focus

of the pilot study. The pilot study and the literature review give insight into why certain students score the way they do on a test that is meant to measure the level of basic knowledge among all students. How does socioeconomic status and primary language, affect the way an English language learner scores on a standardized test?

## THEORETICAL RATIONALE

The theory behind the research on this topic comes from Spring (2005) and his theory of equity in education. Spring explains that political issues have affected the way an immigrant student is educated throughout the history of education. Spring gives a brief history of education beginning with the political history leading up to the No Child Left Behind Act (NCLB) of 2001. Spring explains the reasons behind education as a major platform for politicians. He believes that education is used as a platform to swing a vote even if the politician never intends to follow through with what he promises. Spring breaks down the reign of each president, from Reagan to Bush, Jr., and explains what their ideas for education were. The NCLB Act was presented as an amendment to correct misgivings from Title 1 or the Elementary and Secondary Education Act of 1965. There are titles within the NCLB, namely Titles III and VII, that were meant to deal with cultural inequalities in education for immigrants from certain populations, specifically from populations conquered by the United States. The amendments originally came from the 1965 Elementary and Secondary Education Act. Spring explains that, “The 1965 Elementary and Secondary Education Act was part of the Democratic Party’s so-called War on Poverty. A large and important part of the political constituency of the Democratic Party was composed of groups that had involuntarily become part of the United States as a result of territorial expansion...” (p. 7, 2005)

According to this bylaw, Latinos originally from Mexico would fit into this category because they lost their land in the Mexican-American War as a result of territorial expansion. Spring explores the idea that rather than fixing the problem of inequality in education, NCLB furthers the problem by forcing standardization for all



students and teachers. NCLB expects all learners to be homogenous after claiming to acknowledge differences due to the inequality and suffering of many cultures. Spring's theory brings issues of equity to the surface and analyzes why standardization in education will not set the tone for educational equality.

#### Assumptions

I assume that outside elements are not considered by those who score the STAR and by those who assess the results. I also believe that because of monoculturalistic wording, the STAR is discriminatory. Students who speak the same language yet are from different regions, speak in different dialects or stem from different cultures will most likely read the test questions differently from one another.

#### Background and Need

The need for this type of study is apparent in the research that has already been written regarding similar topics. Past researchers have thought of similar ways to discover how the language used on tests affects test results and if a culturally biased language can skew test results. Some of the key studies that bring these important issues to light are outlined below.

Solano-Flores and Li (2006) conducted a study that explores how generalizing results among students who are considered ELLs can improve the dependability of test results. This theory shows that the language a student should be tested in should be determined by testing the students in their home language and then in the majority language and examining the dependability of their scores. There is more research needed to make this study more accurate. One would need to determine if, when testing students in English, more dependable scores can be acquired if they are tested in their local dialect

of English rather than in Standard English. This would act as a control and prove whether or not dialect in the home language would be a determining factor.

Duncan and Parent, L. D. R. (2005) provide a comprehensive study of the effectiveness of a dual language test booklet when conducting mathematic tests. This study shows the effectiveness of allowing dual language speakers to have the opportunity to see the directions in two languages when reading the directions for taking a test. Further research is needed to show that standardized tests in all subjects would benefit from dual language testing booklets as this study is limited to Mathematics.

Rock and Stenner (2005) suggest further avenues for research and why students of different races test at different levels of school readiness. The authors explore the idea of a construct bias. Additionally they look at the use of biased vocabulary in the tests themselves and recommend that this should be explored in further research.

## REVIEW OF THE LITERATURE

The literature reviewed is divided into categories that keep the material organized and show the evidence of studies that have been done and theories that have already been explored surrounding the topic of standardized testing and the influence of socioeconomic status. The categories are as follows: 1) the effect of socioeconomic status on educational attainment; 2) the relationship between standardized testing and language; and 3) standardized testing and NCLB with an emphasis on the Latino population.

The first section shows that socioeconomic status affects reading in a foreign language as well as a primary language and that there is a direct correlation between the socioeconomic resources of a family and the test score gaps among minorities. This section also focuses on the status of immigrants affecting the education of immigrant children based on educational expectation.

The second section of the literature review covers the successful ways that giving a test with a dual language option has improved the success of language learners. The theory of generalizability is used to show that dialect associated with status can affect the test scores of ELLs.

Lastly, the third section focuses on the correlation of test scores among Mexican American students. The section of the review centers mainly on broader topics associated with standardized testing and NCLB. This section offers some ideas and theories regarding solutions or alternate measures of assessment besides standardized testing. This section presents a successful model of community based assessments and explores the idea of separating children by race and income to satisfy some NCLB requirements. This

section also looks at the controversial idea of having students who traditionally score low on the test, stay away and not test at all. The literature that has been found regarding these topics has all be reviewed separately in order to conduct a thorough examination on socioeconomic status and the ways that it affects reading and standardized testing.

#### Socioeconomic Status and the Attainment of Education

In the first division of the literature, the ways in which socioeconomic status affects reading and the acquisition of knowledge is explored. Kahn-Horwitz, Shiron and Sparks (2006) investigate how skills in one's first language will affect the acquisition of a second language and then compounds that with the effect that socioeconomic status will have on the same language skills, specifically reading. The authors conducted a study in which students from different socioeconomic regions of Israel are grouped based on their reading skills in their first language (L1). These students were of elementary age. They had received signed permission from their parents to participate in the study. They were asked to read words, pronounce them correctly and then asked for the meaning. This determined how well they were able to sound out the word phonetically and if they could comprehend the meaning of the word as well. This was done in their L1 first and then in English to see if there was a correlation between the scores that they received.

Their socioeconomic status (SES) was determined based on the neighborhood in which they lived. The census for the region was divided into categories based on factual information received from the state. They looked at the education level of the majority of people living in that region, annual salary, how many were receiving financial assistance from the government and the median cost of housing. The schools that fell into these

regions were given a numerical score from one (lowest) to ten (highest). The school examined in this study all fell into the categories of four or six.

Admittedly skewed, the authors use this information, although the schools fell so close in range that it was nearly impossible to prove a correlation between SES and L2 language acquisition. The authors used these results as well as other studies to show the correlation here between these two elements. The results showed that students who fell into a low SES category were behind an average of one year in their knowledge of letter sounds and letter names (2006). This study effectively backs up the claim that socioeconomic status and reading acquisition in the L1 can have a negative effect on the student's acquisition of the L2.

The next study examined family status accounting for test score gaps among certain ethnic groups. The theory that SES affects the test score gaps is explored in this study but the issue of race and ethnicity affecting test score gaps come up as well. "What might be causing such gaps? One prominent possibility is that the historical racial and ethnic inequalities in the United States have created disparate socioeconomic circumstances for the families in which white, black, and Hispanic children are reared"(Duncan & Magnuson, 2005, p 36) The authors chose not to ignore the issue of race inequalities while acknowledging socioeconomic inequalities.

The article makes a distinction between socioeconomic status and socioeconomic resources. The distinction is made here in order to put the various resources into categories that can then be analyzed individually. Social dimensions also fall under the category of socioeconomic resources as they do not, necessarily, determine status. These resources include: income, parent education level, family structure, and neighborhood

conditions. They also can include: dropout mothers, single parent homes, multiple siblings, spanking, few children's books, low birth weight, teen mothers, and post natal depression.

A graph within the article shows that, statistically, poor children experience each of these negative social dimensions on a greater scale than children who are not poor. This is determined by examining the data from the Early Childhood Longitudinal Study (ECLS-K). The socioeconomic status of the children used in this nation wide assessment, matched the gaps in test scores. Social hardships, such as the ones mentioned above were measured when the data was taken to determine the status of the children involved in the study. This article presents different view points to demonstrate that, socially speaking, it is not always possible to determine causality. "One should not necessarily infer that eliminating the income gap would eliminate the achievement gap" (2005, p. 39). This point is a common thread throughout the article as the author presents solutions that are plausible but not clearly defined.

Another point of view considers the social dimensions of families whose income fell at the poverty line and compares those dimensions with families with moderate income. The results show that "Children whose families faced deep and persistent poverty fared the worst and registered the largest achievement gap..." (2005, p. 40). This study is then compared with another study which considered a longer list of reasons for the achievement gap found "a considerably smaller difference between low and high income children" (2005, p.40).

The author concluded that although we can see the correlation between socioeconomic resources including social dimensions, we cannot definitively state that

this *causes* the gap in achievement or test scores. Although the research cannot determine a cause, it does offer various ways to try to close the gap. Parental education is one suggestion. “Children with highly educated parents routinely score higher on cognitive and academic achievement tests than children of parents with less education” (2005, p.41).

Tax refunds to decrease poverty are another solution offered to begin to bridge the gap through the leveling of the socioeconomic playing field. Interventions to promote marriage and decrease divorce are social dimension solutions that are offered to begin to close the gap between children coming from single parent homes and those children who grow up with both parents in the house.

The final solution offered is to give families the chance to move from high poverty neighborhoods to low poverty neighborhoods. This option may offer children the opportunity to be exposed to positive peer influences, strong schools, a greatly reduced level of community violence and increased police protection. The theories suggested in this article show that there is a correlation between socioeconomic status and resources, and achievement on tests and school achievement.

Feliciano (2006) explores the effect of socioeconomic status on educational success with an examination into how the pre-migration status of immigrants affects the educational attainment of the second generation immigrant children. This study is conducted using United States census data from 30 immigrant groups and data from the Children of Immigrants Longitudinal Study. The segmented assimilation theory was also used to attain many of the theoretical points measured here. The author used the CILS, which was designed to examine the process of assimilation and adaptation of U.S. born

children with at least one immigrant parent, for her measure of educational status among immigrant groups. She then used this information in her own study using these populations as a sample for an inferential study.

This study showed that if the migrant parents of children being educated in the United States have high expectations for their child's educational attainment, then that child will succeed in their educational ventures. (2006) In order to have high educational expectations for their children, the immigrant group must come from a socioeconomic status that values education.

Many immigrant groups also show the effects of the ways in which their group as a race or a nationality were treated as immigrants to the United States. For example, if the group had a reputation for being elite in their home country, they would be treated as a "high status immigrant group" (2006). Also, if the ethnic group as a whole, valued education in their home country, their status as immigrants would go up because their educational expectations for their children would be high, therefore, their children would become productive members of society.

Many Asian immigrant groups including: Chinese, Vietnamese, and Koreans, have a higher immigrant status than other immigrant groups in the United States. (Feliciano, 2006) She attributes this to the fact that most of these groups have high educational expectations for their children. "Chinese immigrant parents who experienced substantial loss of status after migration had high expectations for their children in the United States partly because they saw their children's success as a reward for their sacrifice and downward mobility" (2006, p. 284).



Although an immigrant group may lose status when coming to the United States, expectations for their children remain at a high level or even increase. These expectations can keep an entire race of people in a high immigrant status category. On the other hand, Hispanic groups, such as Mexicans, are often seen in a more negative light. This may be due to the fact that their socioeconomic status in their home country was low and because of this their expectations for their children's education is low.

A chart included in the article shows that Mexican immigrants have one of the lowest premigration socioeconomic status' and because of this Mexican youths from families with a low premigration SES' have low educational expectations set by their families. "less than 65% expect to attain a college degree" (2006, p. 294).

These numbers show that a families premigration SES, regardless of their post migration SES, can be a determining factor in the educational attainment of their child. Feliciano goes on to make the distinction that certain ethnic groups can not be placed into a category based on the SES of individual families. Some migrant groups must be looked at as a group because that group status can be more of a determining factor than SES. When considering national origin, SES may not have anything to do with the educational attainment of that groups youth.

Certain groups are placed into the bracket of high status immigrants, which leaves some groups placed into a low status category. These categories are often defined by the groups SES in their home country and by the group's educational expectation for their children. One element that must be considered along with the other factors is that certain social and economic resources are offered to groups that are thought of as high status immigrants or, in other words, groups that can assimilate into the American ideal society.

These groups are offered resources outside their homes and families that can help with educational attainment. Because of this, the SES of the family does not have to negatively affect the educational attainment of the child.

The ties to a community are often intensified when a group first begins to settle in a new country. If that ethnic community values education, they will have a positive influence on new immigrant children. They will have access to co-ethnic peer groups that will encourage them and help them form other positive relationships such as those with teachers. If the community that is here is a not a positive influence or does not have the resources to encourage education, than the immigrant child only has the family resources which may be scarce and which may not emphasize education but working instead.

Overall, the article presents the view that a minority group's socioeconomic status in their home country, as premigrants, has an effect on the second generation's educational attainment. The premigratory socioeconomic status of an individual family will not always be the deciding factor in determining the educational expectations of a second generation immigrant child. It is often the case that the social standing of an entire ethnic group in their home country, must be factored in as well. This affects the educational expectations of a group whether they were an elite group in their home country or a group of a traditionally low socioeconomic status.

#### Standardized Testing and the Correlation to Language

The next section of literature explores the correlation between language acquisition, language learning, language translations and testing and test scores. This section covers the literature that has been written and reviews studies that have been done

to show the ways in which 1) language acquisition varies and 2) how the language used on a test can affect the outcome of the test among English language learners.

The literature shows the correlation between language acquisition and test scores in a study that explores two tests; one in English and one in Spanish (Cascallar & Dorans, 2005). These tests are given as college entrance type exams to students in secondary school. The study attempts to find a connection between the two tests and although the two tests are similar, a direct correlation cannot be found because each test, the PAA and the SAT I, was designed for a particular language and it is impossible for certain elements of the test not to get lost in translation. The research team then used the quantitative methods of concordance and scaling to compare scores and establish linkages. The ESLAT test is used to determine bilingual levels and then is used as a screening variable to determine if language is the key factor in proficiency gaps.

Although this study failed to show that these tests could be compared on an exact scale, the study does bring up many pertinent points regarding the correlation between language acquisition and testing. This study shows that although tests do not necessarily translate well because standardized tests are often designed around one language, what it does prove is that a student who takes a test in a language that they are not proficient in, no matter what their levels are like in their home language, will not score as well on that test as they would if the test was translated. This is due to the fact that reading a test question in a language they are not proficient in will inevitably cause them to make an error in decoding or, if nothing else, will cause them a loss of time. "...it is quite possible for a talented second-language student to receive a below-average score on the GRE

Verbal test simply because it takes them longer to read the passage” (Cascallar & Dorans, 2005, p. 341)

Secondary students taking any standardized test will experience the same result. It will take them longer to read the passage unless it is translated into their language of greatest proficiency. Therefore, a test that is not translated will not prove an accurate measure of comprehension ability, all it will show is a lack of language proficiency. This study was conducted as a way to provide a “cross-lingual assessment” (2005) to students to achieve a fair and accurate measure of their ability. The study suggests that to provide a fair assessment, a student should be able to choose which language they feel they would test the strongest in.

Although this study is regarding college entrance level testing, the idea is applicable for all standardized tests with the ultimate goal measuring ability, not English language proficiency. The study investigates language in an even more intricate way by picking apart the fact that tests are designed for English speakers depending on the region where they live while tests are designed for Spanish speakers depending on their language. For example, students from Puerto Rico speak a different dialect than Spanish speakers from Mexico. These languages differ “much in the same way as, say, residents of the United States and residents of Great Britain; both groups speak English, but the nuances of the language differ in the different countries” (2005, p. 339) Item analyses were conducted to account for these differences but many countries are placed into the “Hispanic” category when taking college entrance exams.

To reiterate, this study did not show a direct connection between the two tests, nor did it show that translating test questions and reading passages is the answer, it did show

however, that testing in a language that the student does not have a strong mastery of, is not an accurate way to measure the student's ability. It proves that giving the student a choice of languages would be a more accurate way to measure results. It also showed that there can be dialect and nuance discrepancies when placing testers in a broad category such as language.

In line with the study of language on the effects of test results, an experiment was conducted that addressed the usefulness of a test booklet that translated the test questions into Spanish on one side and left one side in English. The study was skewed based on the fact that the only subject the students were tested on is math which would not necessarily relate to testing on a broader frame. The study also found that testing a student in math in the same language that the student is instructed in, is most often the way that the student will achieve the highest score on a math test. This would make sense as the student will recognize key words in the language that he or she learned the concept in. The second way that the study was skewed is that it used a fairly small sample group and, for the purpose of the study, within that sample group were students who were bilingual.

The dual language test booklet was designed for students with three or fewer years of English language instruction. However, students who were fully bilingual also took the same test. This made the results inaccurate. The students who were proficient in English reported that they rarely used the dual language test booklet, which seems obvious as the students were English proficient. Unlike the last article reviewed, this study allows for the fact that different nationalities within the same linguistic community of Spanish, speak many different dialogues. Considering this, the research team put certain words in parentheses in case they were interpreted in different ways. They also

agreed upon the common for of *usted* instead of *tu* because *usted* is always the polite term whereas *tu* can sometimes be considered rude. Another culturally considerate element to this study was that the research team made it a point to be sure that the translations would seem accurate. They realized that things can be lost in translation (Cascallar & Dorans 2005) so they used qualitative and quantitative ways to make sure this did not happen.

The researchers met with bilingual focus groups to make sure the language was clear and surveyed students in writing about the accuracy of the test. They even added questions to the math test booklet itself to act as a critique of the translations in case the students felt they were inaccurate after taking the test.

Although this study was flawed in some ways the study did show that proficiency in English was the common factor in measuring achievement on the test. This study also showed that the dual language test booklet was not a good tool for everyone. This is not to say that having a standardized test translated into a variety of languages is a hindrance. Additional research must be conducted to determine for whom a dual test booklet would be appropriate.

The use of the Generalizability (G) theory is used in the statistical analysis of test score results among Haitian-Creole speakers. The dialect spoken by a group of people can have an effect on test results even among people who speak the same language. (Solano-Flores, & Li, 2006) This study consisted of students from the same linguistic communities being tested in Haitian-Creole and English. This study also factored in another sample of students that were tested across languages and across dialects as well. These students all spoke in a dialect or code of some kind, as do all people regardless of their linguistic community (2006)

Once these quantitative tests were scored, according to rubrics in English, data was analyzed to determine the statistical significance of the test scores across languages and across different codes. Software for G theory analysis was used once the data had been collected. G theory measured how generalized results were on the generated tests made up for this study.

The results were calculated using formulas generated by the software that multiplied student x rater x item x code. This software left the research team with information necessary to examine the differences between the levels of an individual's knowledge and the effect of code or dialect. The software also showed where measurement error could occur. Different versions of the same test were given to different students who spoke the same language. "There are multiple forms in which an item can be written in a given language. Each form is a unique combination of lexical and grammatical features that are consistent with the rules of that language." (2006, p. 16) In other words, if two people who spoke the same language were asked to translate a test in that language, the translations would not necessarily be identical because of differences in dialect and code. The results showed that code affects the results of a test on a greater scale than language or translating alone. "The results suggest that language as a source of measurement error operates at the level of dialect. Regardless of the language in which ELLs are tested, their performance is sensitive to that dialect of the language in which they are tested" (2006, p. 17)

This being determined, the researchers continued their work to find a way to minimize the degree of error due to code. A second study begins with the creation of

eight new models of testing ranging from monolingual standard and monodialectal standard through to bilingual and bidialectal equivalents (2006).

After these tests were given the results were sorted between codes and languages and results were calculated to determine the “minimum number of items needed to obtain dependable scores for each model.” (2006, p. 17) This study was admittedly skewed in that measurement error due to the rater of the test was not considered and all students who participated in the test spoke only Haitian-Creole; if students who were English language learners who spoke different languages were surveyed the results may have proven different results. What this study does show is that the Generalizability theory can be used in the testing of ELLs by proving that dialect can affect the outcome of the test results of students within the same linguistic communities. (2006)

#### Standardized Testing, No Child Left Behind and Minorities.

The final section of the literature falls under the categories of both standardized assessments and the No Child Left Behind Act. These two issues are a prominent piece of this literature review because they affect assessment and measuring student success. There are several solutions offered as well as close analyses conducted to determine ways that these issues can be ratified.

In the state of Nebraska, an alternative to standardized testing is successful in providing an equitable education for all students. This approach is known as STARS, or School-based Teacher-led Assessment and Reporting System (Roschewski, Isernhagen & Dappen, 2006) This community based program states that each of Nebraska’s school districts is responsible for developing a system to assess student performance. This



system is based on state law required standards but the assessments themselves are created and rated by teachers who have been schooled in the area of assessment literature.

This article was written based on several studies that have been done and includes data taken quantitatively, by comparing percentage scores of the amount of improvement shown per district, and by analyzing the ways in which these districts meet state accountability scores. The results of these studies are printed within the article to show that STARS provides a positive and equitable education for all students. These alternate assessments are based on state required standards and, results show, that this way of aligning assessments to standards is making positive progression in scores of students for each of the 500 school districts in the state (2006).

Advocators for STARS believe that “this process honors teachers and relies on their professional judgment, but it also demands hard work and a great deal of leadership from all of the states educators.” (p. 434, 2006) This process requires that three steps are followed in order to create assessments that meet state required standards, and because of this the educators themselves are doing more work. They must first read and understand all state and locally adopted standards, then they must develop an assessment plan that includes measurable criteria that meets the standards for each grade level. The last step is to put together a portfolio of the assessment procedures and sample assessments so that they can be reviewed and rated by the state. This must be done to ensure that the assessment materials used to measure student learning outcomes are aligned with state standards. Educators in Nebraska are willing to do this because of the equitable and positive results that it has on their students (2006).

STARS is unique because it is a variety of locally formed assessments, not a state test, yet it is considered reputable by the state. Once these portfolios have been submitted and approved they are reviewed by assessment experts from across the nation (2006). These portfolios have to pass through many regulations in order to be considered valid. There are six criteria that the portfolio must meet: The assessment follows the state standards, the students show that they have had the opportunity to learn, the assessment is at the appropriate level of student achievement, the assessment has been scored reliably and the mastery levels have been appropriately set (2006). The results of these portfolio rankings are made public in the states annual *State of the Schools Report*. This is similar to test results being published for the public to see. The people of Nebraska feel satisfied that STARS is working as an equitable and reputable measure for the success of all students (2006). The ratings that are published show that each district is meeting up to expectations of the state by placing rankings in categories of Exemplary, Very Good, Good, Needs Improvement, and Unacceptable. In order to meet accountability goals set by the state of Nebraska, the schools must rank at Good, Very Good or Exemplary; if they don't they must develop and improvement plan to improve in three years or face loss of accreditation for the district.

Each year results rise, percentage wise, and do not drop. "In 2001, 66.31% of Nebraska districts had already earned ratings of good or better. By 2005, the percentage of districts meeting the goals had risen to 97.73 % in reading and 99.54 % in math." (2006, p. 435) STARS works because teachers and educators develop assessments that honor different learning styles to measure success. This takes teachers away from teaching to the test and takes the student out of the high risk testing scenario.

Integrating schools by socioeconomic status may even out test scores and will fulfill the ideals of NCLB by providing equal opportunities for education to all children regardless of race and socioeconomic status (Hardy, 2006) A model school, Washington Gifted and Talented Magnet Elementary School, in Raleigh, North Carolina is reviewed and examined due to the fact that it has been considered a National Magnet School of Excellence for four years and it scores high on state standardized tests despite its population of students who fall into the category of low socioeconomic status. “32% of students receive free and reduced-price lunches.” (Hardy, 2006, p. 2) This school is a neighborhood overlooking a housing project development yet parents want their children there for the diversity and the outcome of academic excellence they provide.

This program has attracted national attention for its ability to make the school diverse in a way that does not make race-assignment distinctions. Meaning it does not pull children of different races from low-income neighborhoods just to meet desegregation requirements. This school is referenced often in the article reviewed here and used as a model for solution because of its high success rate for all students. Economic integration and racial desegregation are two different things. The model school in this article uses economic integration by allowing for equal opportunity while desegregation is considered to be assigning students of different races to schools outside their neighborhoods with the assumptions that race is tied to socioeconomic status.

Districts are beginning to prohibit racial desegregation due to the fact that it is considered a social engineering issue where as economic integration shows major academic advantages for low income students. Being in schools that have college preparatory classes and mingling with students, who have high parental expectations at

home, make these schools places where low and middle/high income students start off on a level playing field. “Studies and test scores show that students do markedly better in middle-class school, for perhaps self-evident reasons: better teachers, stronger discipline, more college prep courses, and peers who believe from an early age they are destined for college.” (2006, p. 2) This trend seems to be successful in providing opportunities for students while staying accountable to state standards.

For those proponents of integration who claim it is a social issue, this article proves that it is very much a social issue and that integration promotes social cohesion as well as equal opportunity (Hardy, 2006). The fact that there are proponents for this issue, who want schools to go back to being separate in terms of neighborhoods, brings up the ironies about this topic.

The first thing that the article classifies as ironic is that public schools are getting more diverse as the population gets more diverse. Public schools nationally are only serving 58% white students which is a drastic change from the 1960s where 4 out of 5 students were white (2006). This number increases every year and is expected to keep increasing.

The second irony occurs because as schools begin to be considered successfully desegregated, meaning they have attempted to eliminate racial segregation, they begin to stop assigning students to schools and just allow them to attend neighborhood schools. Because of socioeconomic status and housing patterns, this brings back segregation at an even more extreme level.

The solution outlined in the article would steer away from racial desegregation and toward economic integration. This solution is to integrate housing instead of school.

The integration of housing and neighborhoods would take the racial segregation out of neighborhood schools as well as eliminate the “bussing” of low income or urban students into middle income or suburban schools. This would also fulfill the goals of social cohesion and put students into social networking scenarios with students whose families had high educational aspirations and goals. This model has succeeded in Vancouver Canada and was recently attempted in San Francisco, when mayor Newsom proposed 20% of new housing be designated not as low income but as ‘family friendly’ (2006). Economic integration seems to be a valid solution to the issue of providing equal opportunities to children as other solutions have not proven to be successful. For example, attempting to address achievement problems in low income schools by raising standards or reducing class sizes has not worked. These attempts have done nothing to raise test scores among students considered to be low socioeconomic students.

“Economic integration is a good policy, both for its potential to raise achievement of low income students and to create the kind of integrated schools vital to American democracy. Integrating schools economically succeeds in celebrating diversity and has shown an increase in academic achievement for students who may not have succeeded elsewhere.” (2006, p. 8)

The article concludes by revisiting the model school and outlining exactly how this successful implementation works. “The district maintains what it calls a Healthy Schools Review Team that examines each school in depth to determine its challenges, such as students performing below grade level, and devises specific plans to address its needs. For example, one schools plan may include hiring teacher’s aides or expanding tutoring programs” (2006,p. 8). The solution of economic integration for schools and

neighborhoods is a solution that has worked to implement NCLB in that each child is provided with an equitable opportunity for education.

Providing students with an equal opportunity proves to be quite challenging at times and the NCLB act can lead schools astray in their search for equal opportunity. Some schools will go as far as to manipulate test results because of the high stakes involved with NCLB. (Lemke, Hoerander, & McMahon, 2006). A study that examines the ways in which schools manipulate test results to achieve the 60% passing status required by NCLB looks at the variation in test results as the scores of students who did not take the test receive a score of unknown. This score is not factored into the overall results so the school receives a higher overall score than they would have if those students who did not take the test had taken the test and failed.

This study uses a bounding technique to show that classifying schools as passing or failing against a fixed percentage can often be misleading. This research also uses regression analysis to show that schools might be working the system. This regression analysis looked at the past scores of certain schools who were on the verge of not passing and then at the number of students who did not take the test or who took a modified version of the test. If the amount of non-test takers increased and the over all score of the school increased, the researchers made the inference that some manipulation had occurred.

They included charts showing the standard deviation as well as the percentage of possibility of error to show that there is a statistical significance between the increase in test scores in some school districts and the number of non-test takers. NCLB only requires that 95% of a schools population take the test and that special populations

(economically disadvantaged, major racial and ethnic groups, disabled, and students with limited English proficiency) be considered separately. This percentage serves as a loop hole for some schools to eliminate certain students from taking the test. Statistics show that not only do these poor academic students not bring down the test scores by taking the test but their absence can actually be the difference between the school becoming a program improvement school or not. The NCLB has such high stakes for educators that schools have motivation to encourage certain students to not take the test. If a school becomes a program improvement school, they will lose funding and students could receive vouchers to attend other schools.

There are a number of ways that educators can manipulate test scores and eliminating test takers is just one of those ways. The others include: holding students back a grade to postpone testing, relaxing the rules about dropping out of high school, assigning borderline students as learning disabled so they will not take the primary assessment, substituting a well rounded curriculum for one that teaches to the test, or just providing students with questions and answers prior to the test (2006). The data is limited in that it only uses data from 11<sup>th</sup> grade students within the Illinois school district but it does consistently show that with 573 high schools, scores of unknown were changing the results of the testing overall and that schools were using it to their advantage.

The researchers make two assumptions in order to come up with a concise result. They show the results that would have occurred if all the students marked as unknown would have passed the test, then they should the results of what percentages would have looked like if all non-test takers would have failed. The results showed a greater statistical significance that schools would not have met the 60% proficient mark if all

non-test takers would have failed the test. Some schools within the study reported that 100% of students took the test as a result of administrators accurately perusing the students who missed the test to make it up. The research team involved in this study looked at how much higher these schools results would have been if they would have encouraged 5% of their worst students to stay home, and those results showed a much higher percent of proficient scores. These results were used to further prove the theory that some schools whose results were significantly higher than expected were excluding certain students on purpose from taking the test.

Within the Chicago school district, schools were lowering regulations to control which students dropped out of school and were arguably making it easier to drop out. These non-academic students would then not take the test and test scores would remain high (2006). This seems to be happening because educators, including districts and administrators, are now held accountable for these results. It is not as if standardized testing is a new idea, but that level of accountability at the administrative level was not there before. The purpose of this study seems to be to show that overall, the results of the regression study show that schools are selecting students who do not take the test and that having fewer students take the test in the first place, is associated with a higher pass rate. A solution for this problem is offered here in this research and that solution would be to give every student who did not attempt the test, a failing score instead of not factoring the score in at all. This would eliminate the practice of purposely not including students and would force the administration to try to have every child take the test. This would show a reading more accurately because all of the student's levels would be taken into account.



The idea of test score gaps among students from different backgrounds starts at an early age. The gaps are apparent as early on as kindergarten and seem to follow the same patterns. The gaps occur between white students and black or Hispanic students. These gaps are not racial however; they have the highest correlation between students of varied socioeconomic status. Race correlates to socioeconomic status on many studies that have been conducted and that is where studies often make the connection to racial gaps.

Rock and Stenner (2005) examine common tests that are given to students to gauge kindergarten readiness and then chart them according to students race, outside factors (socioeconomic status), and the reliability of the test. The Peabody Picture Vocabulary Test-Revised (PPVT-R) is a test with a high reliability that often shows a large gap between the readiness of students to begin kindergarten. These results were acquired by examining the National Longitudinal Study that tracked a sample of 12, 686 young people in 1979 and then interviewed them each year through 1994 (2005).

Another large sample was taken from a study known as the Infant Health and Development Program. These 985 infants were born with low birth weights and their development was tracked from 1985-2000 (2005). These children were born in eight different cities and their development was tracked using several tests including the PPVT-R. Using these sample populations, the PPVT-R found marked differences between black and white readiness for kindergarten. This gap becomes known as the “raw gap” (2005, p. 25). Once outside elements are factored in, scores are adjusted to take into account: age or education of a child’s mother, family income, or whether the child was born with a low birth weight.

This research is not completely sound in that the PPVT-R often finds a larger gap between white and black students than other tests do and once other factors are considered the standard deviation rises to more than 1, while the raw gap stays at .05. This lowers the reliability of the test itself because the results are not always consistent. This adjusted gap show that outside elements do have a substantial impact on how prepared a child is for school. “This pattern suggests that influences outside school, such as family background, health and neighborhood can have important effects on a child’s academic readiness for school.” (2005, p. 27)

The last issue addressed in this analysis is whether there is a racial or ethnic bias present within these readiness tests that are setting certain groups of people up for failure. It is suggested that this gap may be caused by a systematic bias in the test. “Perhaps certain vocabulary words are more commonly used in white families than in black or Hispanic ones” (2005, p. 28) A way to figure out the bias is to look at the results from groups of white and non-white students who have the same score on the test. They should have the same questions wrong if the test truly tested ability. If the students have different questions wrong based on the wording of the questions themselves, then “differential item functioning” (2005, p.28) exists. Another bias that may be present in standardized tests made up by people from normative backgrounds could be construct bias. Perhaps it is bias in the test or outside factors that create this gap but what ever it may be, a solution must be found. Although new methods to measure readiness are needed to learn why the gap is there, it is undeniable that a gap exists and, after all the research is done, it is still a perplexing issue.

### Summary of Major Findings

In summary, the points found throughout this review of the literature suggest that socioeconomic and social status do affect the way a child takes a test. The literature suggests that this is due to a variety of reasons: the parent's expectations of the child's education, the dialect in which a child speaks will affect the way they comprehend the language on the test, and the flawed system of standardized testing in general, to name a few. Each study that was conducted seems to prove that a number of outside elements and factors are going to affect the way a child takes a test. Some children are not going to be able to read the test questions given in English only, and some students who can read English are not going to understand the language on the test because no one speaks to them at home using high level vocabulary words. Both of these reasons have to do with socioeconomic status. Even the status that an immigrant group had before they migrated, will affect the way a child takes a test. If an immigrant group is not thought of as a highly assimilated immigrant group, the child may react to the test differently than other children.

Another reason for the gap in test scores among English language learners is the standardized tests themselves. If there was an alternate way to assess students, such as a portfolio of student work, a more accurate measure of student knowledge could be attained. The No Child Left Behind Act has left educators with the feeling that standardized tests are the determining factors in how the success of their students, and their own success, will be measured. As teachers, administrators and school districts feel the stress of standardized assessment, they begin to resort to measures that are not necessarily beneficial to the student.

According to the research, things are being done that actually harm student learning instead of encourage it and can even go as far as to discriminate against certain students. Students who do not speak English proficiently might be encouraged not to take the test or, if they miss the primary assessment, they might not be instructed to retake the test. This is another detrimental factor that leads to the poor test scores of English language learners. These students might even be encouraged to drop out of school so that their test scores do not reflect on their school district. Teachers also will often only teach to the test so that the scores go up and will leave out other important content. This will put many students at a disadvantage as they move ahead in grade levels without the necessary background knowledge. Standardized testing in general is changing the culture within some school districts and is doing nothing for the growth of English language learners.

#### Limitations and Gaps in the Research

Although the literature that has been reviewed thus far brings about many valid points, there is still room for additional research. Many of the studies that have been conducted admit to using a small sample size or a very specific sample. For example, the sample might only use English language learners who come from one linguistic community originally or standardized tests from one region within the United States. Because of this, many of the studies have larger margins for error as they cannot infer their results on to the population as a whole. In my opinion, more research should also be conducted as to what can be done about this problem. The problem being that culturally a discriminative curriculum is causing the gap in test scores among language learners.

Now that this has been documented, what can be done to solve this problem? In two of the articles reviewed, there were solutions offered. One was to translate test language and the other was to create student-centered portfolio that showcased the students work as a measure of knowledge. The first solution was proven down in the study because it did not raise the test scores of the majority of students who used the translated documents. The second solution would take more time and money but its success has been proven state wide in Nebraska.

Another gap existed in the research when results were given for studies that stated that assessments were constructively biased because of the language they used and that the language often discriminated against the dialects and status of non-normative classes. There seemed to be a gap because although this had been proven many times over, a resolution was never suggested.

#### Implications for Future Research

In order to eventually change the flawed system of the standardized testing movement, more research must be done to prove that alternate assessments must be used in order to accurately achieve results. If more research were done about the lack of students that are actually succeeding on the test and the standardized tests themselves, more people would see a need for a solution to this problem and that there is a viable solution in sight. As the population of English language learners' increases and the gap in test scores gets wider, standardized testing will show less accurate results for school districts as a whole. As the scores drop per district and district officials discriminate against who can take the test and who cannot, standardized testing will become less popular and it will be discovered that it is not serving the needs of the majority. When

this happens, solutions must be in place to measure the skill attainment of students from non-normative backgrounds.

Further research could be conducted to be able to prove why solutions are necessary and it could also look into how language could be revised to be pluralistic. Pluralistic language does not necessarily mean that the test would be translated into many languages; it means that the test would use a language that would appeal to many classes of people. It would not be language that would appeal to the normative cultural and it would not discriminate in terms of socioeconomic or social status. Research has proven that standardized tests are often biased and culturally discriminative, so further research must be done about how this language construct can be revamped.

#### Overall Significance of the Literature

Overall, the literature reviewed about the topic of test scores among English language learners and the gaps that exist in these assessments, show that socioeconomic status does affect the way that a student will take a test. Other outside factors affect the score as well such as the social status of an entire immigrant group and the language that is used on a standardized test. If the test is written by someone from a normative background, for someone from a normative background, many people will be left out of that scenario. The research has also shown the flawed system that surrounds the standardized tests. The large impact that state law puts on the results of standardized tests has driven educators to be selfish in attaining ideal test results. This will most often affect the language learner in a negative way because they are, most likely, achieving low scores on these tests. Because a district's worth is often determined by test scores, with out any factors considered, these students can be discriminated against and treated

unjustly when it comes time to test. The research has also shown that very few viable solutions have been offered for this problem. It has been proven that a low socioeconomic status or a low social status standing can affect the way a student does on a test, but it has not been proven how we can move away from the negative and strive for the positive.

If educators know that English language learners are not scoring well on standardized tests, and they know the reasons why their scores are low, then why aren't they fighting for solutions to this problem like alternate assessments? In order to make the justified argument that standardized tests are not a fair or accurate measure of student ability, there must be a feasible solution waiting in the wings.

## PILOT STUDY

### Method

#### *Sample and Site*

One middle school teacher was the subject of the interview conducted for the pilot study. The study dealt with students attending middle school in one school district. This district includes many immigrant students who are learning English as their second language. The interview used for the study was conducted on the school premises where the subject of the interview teaches and where I teach as her colleague. This school site is also where the plan for intervention aimed at ELLs was first implemented approximately five years ago. I obtained permission to conduct this interview by submitting a letter of permission to the principal and the teacher herself. The interview was conducted in the teacher's classroom after school hours. She was given a paper copy of the interview questions so she could look over them as I read them aloud to her. This was done so she could think about and format her response to each question. As she spoke her answers aloud I typed them on a laptop computer.

#### *Access and Permissions*

I obtained permission from the teacher who participated in the interview before the interview took place. She signed a consent form approved by the Institutional Review Board and was given background information on the study. I also explained verbally what the study was about and why I thought she would be an ideal candidate for an interview. I had the opportunity to work closely with this person as a colleague so it was convenient to speak with her about this study frequently. I also submitted the necessary letter to the principal for whom we both work to make sure she was aware of the topic of my study and the fact that I would be interviewing a teacher on school premises. During



the interview the teacher was informed by me that she could pass on answering any questions that she felt uncomfortable answering and we could end the interview at any time. The interview subject was very cooperative and helpful in answering all of my questions.

#### *Data Gathering Strategies*

I used an interview with ten questions to gather data for this study. The interview was conducted in an informal setting where the interview subject was comfortable. I formatted the interview questions so that optimal information could be gathered based on the subjects knowledge and expertise.

#### *Data Analysis Approach*

The data gathered for this study was analyzed using qualitative analysis. The information gathered during this interview was gathered in a question-answer format. The data was taken through dictation and then each verbal response was compared to the literature found during the research process. Similar themes were found and comparisons could be made between the responses gathered during the interview and the research that was gathered pertaining to the subject. The data gathered during the interview made it possible to format this pilot study because hypotheses could be made in relation to the literature and the findings from the qualitative questions used in the interview.

#### *Ethical Standards*

This study adheres to Ethical Standards in Human Subjects Research of the American Psychological Association (Publication Manual of the American Psychological Association, 2007). Additionally, the project was reviewed and approved by the Dominican University of California Institutional Review Board, Number 6066.

*Results or Findings*

The results of the pilot study connected thematically to the literature on the subject of ELLs and standardized testing. The subject interviewed for the pilot study does not believe that there will ever be a test that is capable of successfully measuring the knowledge and language acquisition of language learners. Her idea, in response to this question, is regardless of the race or culture of a person, if the person is from any type of non-normative background, they will have less of a chance to excel on a standardized test. The literature reviewed in the research component of this study suggests the same thing.

The pilot study and the literature are also linked when it comes to the way that socioeconomic status of the parents affects the test scores and education of a student. The interview subject explains that if a parent does not expect their student to do well and does not take an active interest in their education, the student does not stand a chance in school. The literature suggests that if a parent's educational expectations are low because of cultural norms, this will affect the way a child does on a test or in school in general.

The expert interviewed also had a theory of why the "*passports*" program did not work within the school district where it was piloted. She explained this as a cultural phenomenon. When *passports* was first presented it seemed like grouping students together based on language acquisition would help them to excel in school because they would not feel like they could not do the same work that other students were doing. What this really did was perpetuate the isolation of the Latino students. They became segregated and this made it harder for them to branch out. Instead of only being together in the classroom they began to only socialize with each other and they never met any

other students at all. This was detrimental to them and in some cases it encouraged the gang culture that has been ever increasing in this school district. This grouping or segregation also made it easy for these students to brush off low test scores and failing grades because they were all seeing the same thing. This compared with a culturally low educational expectation, took the shame out of doing poorly in school. This information was attained during the interview but the literature that came about during the research phase is there to back it up. Due to this backfire of sorts and major district budget cuts, this plan for intervention is not being continued next year. There will be no plan to try to bring up the standardized test scores among language learners.

#### *Analysis of Themes*

Through an analysis of the pilot study and the literature, many common themes can be found. It is not a new idea that self-segregation within the Latino culture can be tied to lower standardized test scores. The ideas gathered during the interview present this theory in a logical way.

The interview subject explains that through observation she has noticed that these students are not a positive influence on each other. Instead they make it seem acceptable or even socially redeemable to do poorly in school and on tests. Although this theory has a limited scope, it seems feasible that this is happening to immigrants and language learners in other districts.

The plan for intervention that was in place in this particular school district was a model for success for a short time. The expert involved in the interview suggested a new plan for intervention that involves a more main stream approach. It would be interesting to see if being mainstreamed and pulled out of isolation would force this group of

students to try harder in school. The literature on the subject varies on this topic but one study does suggest that bussing students from low income districts into higher income districts would give them a stronger chance to obtain an excellent education. This idea ties in directly to results obtained during the pilot study. The students would see others around them excelling in school and this would increase their educational expectation for themselves.

### Summary of Major Findings

#### *Comparison of Findings to Existing Literature*

The pilot study conducted correlates to the literature by further linking the ties between socioeconomic status and success in education for school-age students. A study conducted by Duncan and Magnusson (2005) concludes that although the gap in test scores correlates statistically with students whose parents suffer from severe poverty, the study could not prove that poverty was the cause of low test scores. This study also takes ethnicity into account. This can be compared with the pilot study because ethnicity and culture were used as a determining factor to analyze educational attainment in a standardized classroom.

Feliciano's study (2006) compares the educational expectations levels among different immigrant groups. The results of this study coincide directly with the interview that was conducted for the pilot study. Feliciano explains that Chinese immigrants, for example, often come to the United States and lose the status they had in their home country based on monetary value but the expectation they have for their children's educational level remains the same. On the other hand, Mexican immigrants often have a low educational expectation culturally in their home country and it remains low when

they immigrate. Feliciano's theory proves why the passports program for intervention catered mainly toward Latino students. Student's who came into the district from countries with a high educational expectation level, were expected by their parents to make it in regular education classes. These parents opted out of language intervention classes and the students were main-streamed.

#### *Limitations of the Study*

This study is limited because it is based on a pilot study that stems from one interview. This interview was conducted with an expert on English language intervention but it could also be considered one person's opinion. Another limitation would be that the scores and students, who were analyzed and referred to, are students from one school district who were involved in a program dealing with language intervention. There are other students who fit into the category of "language learner" that are not participants in the *passports* program so therefore they were not considered in the pilot study. If they had been, the results of the study might have been different. On a similar note, if more than one expert in the field of language intervention had been questioned, the results of the pilot study could have been altered.

#### *Implications for Future Research*

The scope is limited in this study due the small sample size of the experts that were interviewed and the students that were looked at were from a similar cultural background. Future research could be conducted by interviewing other experts in the field on intervention programs for language learners. This might bring up an entirely different theory of why language learners are scoring poorly on standardized tests. It

would also be extremely informative to do a qualitative survey where the students themselves were able to express their opinions on the issue. It would be fascinating to see the data from the survey that asked them the questions and made them feel that they had control over their own education.

Another possible pilot study would be to interview an expert on the Latino culture and how it is changing as immigration to the United States is increasing. This study would provide researchers with authentic information on the reasons why schools in the United States are failing some of these students.

#### *Overall Significance of the Study*

The significance of the study and the literature that was reviewed prove that socioeconomic status and cultural norms do affect the way that a student is educated and the way a particular group of students score on standardized tests. The focus students in this particular study were from the Latino culture and had begun learning English as their second language. This group of students were limited in scope because they all came from a similar background culturally. This cultural similarity was looked at as an isolating factor within the research and in the pilot study. The fact that the Latino students in a suburban school district were self-segregating and being segregated in terms of class grouping, led to a low academic performance within the isolated group. According to the research and the study overall, this low level of academic performance was due to low educational expectations as a culture and lack of academic role modeling at home. A low socioeconomic status furthered the gap between English language-learning immigrants and other students at the school. The study also showed that the plan for intervention known as passports, which was designed to help the ELL students, did the opposite after

a number of years and furthered the segregation effect. After reviewing the literature and conducting the pilot study, it is clear that a plan for intervention must be implemented for Latino students in order for them to excel in what has become a standardized curriculum.

## CONCLUSIONS

The theory behind the research is that other elements must be considered intersectionally in order to achieve a fair measurement of skills on a standardized test. Theoretically, students from different backgrounds are going to read and respond to questions in a different manner. Different people are going to associate different connotations with each word or phrase that they read. Therefore, each person is going to score differently on a standardized test depending on how they perceived the wording in each question. If the test is culturally discriminative, those from minority cultures become set up for failure beyond the obvious language discrimination. It is assumed that a person who does not speak English is not going to excel on a test given in only English. However, it is not so widely known that if the person who does not speak English fluently also comes from a family with a low level of education in their home language and they lived in a place where a non-normative dialect was spoken due to their status they will be at an even greater disadvantage when it comes to deciphering the wording on a standardized test.

This would fall under the theory of equality in education. The biases present within the standardized testing movement do not present each student with an equal chance at an education. Another theory that could prove this to be true is the Generalizability Theory. This theory would be used to measure the amount of assessment error due to code (language or dialect). This important theoretical study will show that various factors contribute to measurement among test takers. If this theory that students are scoring low on standardized tests because of culturally discriminative language is true, it will be determined through looking at statistical results quantitatively. This will allow for an understanding of the students language acquisition and may also give a first



hand understanding of the students' status and dialect. After an understanding of the students and their various levels of language, a generalization could be made about their test results compared to the level of language they seem to have acquired.

The literature that has been researched and reviewed along with the pilot study can be used to support the concept that cultural and socioeconomic influences affect the way that students score on standardized tests. The pilot study also supports the belief that culture and ethnicity should be considered when analyzing test results because they can have an effect on the way a student goes about taking a test and therefore can influence the outcome or score of the test.

From the comparison of the pilot study and the literature, comes the idea that students from a Latino cultural background can sometime isolate themselves culturally. This can lead to a group of students that teach each other that it is acceptable to do poorly on tests and in school. The pilot study and the literature also agree that parental expectations have a great deal of influence on student achievement and that students who come from extreme poverty can be years behind students from moderate income families.

Overall, after the literature has been compared to the results of the pilot study, it can be concluded that a plan for intervention is what is needed to improve the scores of Latino English language learners on standardized tests. The curriculum of the present day is standardized and it seems as though it will remain that way for quite some time due to political agenda such as No Child Left Behind. The plan for intervention should not further isolate students stemming from one culture but should expose them to students from moderate income families with high educational expectations. This will eliminate the idea that it is socially acceptable to fail tests or even entire grades. The literature

shows that with each generation born in the United States, the income level will rise and, in turn, the educational expectation level will rise. This will help later generations of Latino students to break culture barriers and stereotypes as they begin to excel academically.

REFERENCES

- Cascallar, A., & Dorans, N. (2005). Linking Scores From Test of Similar Content Given in Different Languages: An Illustration Involving Methodological Alternatives. *International Journal of Testing*, 5 (4), 337-356. Retrieved September 23, 2007, from Education Research Complete database.
- Duncan, G. J. & Magnuson, K. A. (2005, Spring). Can Family Socioeconomic Resources Account for Racial and Ethnic Test Score Gaps? School Readiness: Closing Racial and Ethnic Gaps. *The Future of Children*, 15 (1), 35-54
- Duncan, T.G., Parent, L. D. R. (2005). Study of a Dual-Language Test Booklet in Eighth Grade Mathematics. *Applied Measurement in Education*, 18 (2), 129-161. Retrieved September 29, 2007, from Education Research Complete database.
- Feliciano, C. (2006, October). Beyond the Family: The Influence of Premigration Group Status on the Educational Expectations of Immigrants' Children. *Sociology of Education*, 79 (4), 281-303.
- Hardy, L. (2006, September). Separate Our Students by Race and Income to Meet NCLB? *Education Digest*, 72 (1), 12-20
- Johnson, A. (2006, September). No Child Left Behind. *Clearing House*, 80 (1), 34-36.
- Kahn-Horwitz, J, Shimron, Sparks, R.L. (2006). Weak and Strong Novice Readers of English as a Foreign Language: Effects of First Language and Socioeconomic Status. *Annals of Dyslexia*, 56 (1)
- Lemke, R., Hoerander, C., & McMahon, R. (2006, June). Student Assessments, Non-Test Takers, and School Accountability. *Education Economics*, 14 (2), 235-250. Retrieved September 23, 2007, from Education Research Complete database.

- Morales, M. Cristina, Saenz, Rogelio. (2007). Correlates of Mexican American Students' Standardized Test Scores: An Integrated Model Approach. *Hispanic Journal of Behavioral Sciences*, 29, 349-365. Retrieved October 30, 2007 from Education Research Complete database.
- Rock, D.A. & Stenner, J. (2005) Assessment Issues in the Testing of Children at School Entry. School Readiness: Closing Racial and Ethnic Gaps. *The Future of Children*, 15 (1) 15-35
- Roschewski, P., Isernhagen, J. & Dappen, L. (2006, February). Nebraska STARS: Achieving Results. *Phi Delta Kappan*, 87(6), 433-437. Retrieved September 23, 2007, from Education Research Complete database.
- Solano-Flores, G., & Li, M. (2006, March). The Use of Generalizability Theory in the Testing of Linguistic Minorities. *Educational Measurement: Issues & Practice*, Retrieved September 23, 2007, from Education Research Complete database.
- Spring, J.H. (2005). *Conflict of Interest: The Politics of American Education*. New York: McGraw-Hill.