LONG-TERM EFFECTS OF NATIVE HAWAIIAN STUDENTS' EARLY ACADEMIC ACHIEVEMENT UNDER THE NO CHILD LEFT BEHIND LEGISLATION: A MULTILEVEL COHORT ANALYSIS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

EDUCATIONAL PSYCHOLOGY

May 2011

By

J Malkeet Singh

Dissertation Committee:

Shuqiang Zhang, Chairperson
Paul Brandon
Marie Iding
Seongah Im
Curtis Ho

ACKNOWLEDGEMENTS

I would like to thank the members of my dissertation committee for taking the time out of their busy schedules to provide me with helpful comments, moral support and encouragement during the dissertation process.

I would also like to express my gratitude and appreciation to Dr. Zhang, my committee chairperson, for guiding me throughout the dissertation process. He has read and commented on many of my incomplete drafts and provided me with innumerable insightful and helpful suggestions.

ABSTRACT

The focus of the No Child Left Behind (NCLB) Legislation is to close the achievement gaps due to disadvantages based on minority status, socio-economic status, special education (SPED) or Limited English Proficiency (LEP). Poverty and culture have been consistently reported to have an impact on academic achievement. However, there have been few cohort studies that have investigated the impact of early academic achievement on long-term academic success in conjunction with the effects of poverty and culture. Furthermore, no multilevel studies have been conducted to study the impact of early academic achievement on future success from elementary to high school within the NCLB context. This oversight has inadvertently directed attention away from the impact of students' performance at early grades on their future academic achievement.

Among all ethnic groups in Hawaii, the Native Hawaiian student population has the lowest academic performance in Hawaii's public schools. Cultural and socio-economic disadvantages are usually associated with low performing groups. However, the disadvantage of having low early academic achievement has yet to receive adequate attention. Establishing the unique disadvantage of low early academic achievement beyond the disadvantages due to culture or poverty is crucial since early academic achievement may be one important factor affecting the student's future academic performance. A careful examination of the impact of early success on future academic achievement for the 2002 Native Hawaiian cohort was therefore conducted with the White peers serving as the control group.

This multilevel cohort analysis revealed a significant and dominant impact of the academic performance at Grade 3 on the reading or math performance at the fifth, eighth and tenth grades over and beyond the effect of culture and poverty. This impact remained stable from elementary to middle school and from elementary to high school. The current study also revealed that Hawaiian ancestry translates into an additional unique disadvantage on academic performance at the fifth, eighth or tenth grade. This disadvantage increases from the third grade onwards to the tenth grade with early academic performance and poverty statistically controlled for. In contrast, the impact of low socio-economic status remained stable from the third to the tenth grade. Those results were stable whether or not SPED students were included in the analysis.

The findings suggest a need to focus interventions on foundational academic preparation at the early grades. Educators in public schools should also direct more attention toward Native Hawaiian students. NCLB's focus on closing the achievement gaps between advantaged and disadvantaged groups at the school level need to be broadened to allow more instructional attention to be directed towards earlier grades.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION TO THE STUDY

The 2001 No Child Left Behind (NCLB) legislation in the United States was enacted to eliminate achievement gaps in public education. The objective of NCLB is to hold all schools accountable to this goal. Because of NCLB, states are now required to develop standards-based accountability systems to enable all students to attain proficiency in reading and math by 2014 regardless of socio-economic, ethnic and disability status, or limited English proficiency (LEP).

To comply with NCLB, each state has to set its own standards and yearly goals for each school to meet and thus each state has its own definition of proficiency levels. All public schools in the United States (U.S.) are expected to meet the annual goals that are set by their respective states. Schools are held to the same expectation of meeting the state's goals regardless of their initial academic standings. Schools are required to focus on improving minority and economically disadvantaged students' academic proficiency levels, thus specifically targeting student groups that have traditionally performed at a lower level. The main goal is to close the achievement gaps between White and other ethnic groups, such as Black and Hispanic students, and also between economically disadvantaged and economically advantaged groups.

NCLB does not consider a student's initial level of proficiency or a school's initial academic standing when applying sanctions to schools that fail to meet the state's yearly goals. Therefore, NCLB mandates that schools and students who initially performed at a lower level improve at a faster rate than schools and students who initially performed better. In other words, schools and students are expected to improve at different rates to meet the state's goals.

In 2002, the state of Hawaii began to implement the Hawaii State Assessment (HSA) to comply with the NCLB mandate. All schools in Hawaii are required to consistently increase the annual percentage of students meeting the proficiency levels by fixed amounts set by the state in both reading and math (and additionally in science starting from 2008).

Under NCLB, Native Hawaiian students are included in the Asian and Pacific Islander group and not considered a distinct group. Just as Native American students on the mainland United States are among the most at-risk, Native Hawaiians have traditionally underperformed on standardized tests for both reading and math in Hawaii. Academically, they are ranked lowest, behind all other major ethnic groups in Hawaii, thus making them the most at-risk ethnic group in Hawaii (Kanaʻiaupuni & Koren, 2003). One unintended consequence of classifying Native Hawaiians with students of Asian ancestry for NCLB purposes is that the achievement gap between White and Native Hawaiian students is overlooked. Treating Native Hawaiians as a distinct group would rectify this oversight.

While ethnicity is considered an important between-group factor under NCLB, there is also considerable variation among the students within each ethnic group. This variation at the individual level has been left largely unresearched and unaddressed in the NCLB context. NCLB has so far focused on a student's minority status, economic disadvantage, special education and limited English proficiency (LEP) but has not directed adequate attention to early academic deficiency as a causal factor in itself. This has led to policies that are focused primarily on socio-economic and ethnic disadvantages, inadvertently diverting attention from the consequences of early academic deficiency. In other words, how academically at-risk students' low starting level of proficiency affects future academic success has not been adequately addressed in conjunction with socio-economic and ethnic disadvantages. Even though cultural identity and socio-economic disadvantage have substantial effects on early academic performance, they do not account entirely for the individual variation in early academic performance. Thus, the effect due to early academic achievement may need to be considered as a unique causal factor over and beyond cultural identity and socio-economic disadvantage.

NCLB has been in effect for more than half of its life span since 2002. After an extensive literature search, it has been found that there have been very few studies that looked into the impact of early success within the NCLB context. In Hawaii, so far two studies (Takanishi, 2005; Uyeno & Zhang, 2007) have provided evidence of the impact of early academic success on later academic achievement. However, the first study's span was within elementary education using a multilevel approach (Takanishi, 2005) while the second study's span was from the fifth to seventh grade using logistic regression analysis (Uyeno & Zhang, 2007). The current study extended the span from the third grade up to high school using hierarchical linear modeling (HLM). Therefore, this dissertation may well be the first study within the NCLB context to use the multilevel approach to track a Native Hawaiian cohort from the third grade up to the tenth grade, and also the first study to investigate the impact of Native Hawaiian students' early success on their future academic achievement through high school over an eight-year period (2002-2009).

This dissertation addresses the lack of research in the following four areas:

(1) a lack of research on academic achievement of the Native Hawaiian students as a distinct group under NCLB,

(2) a lack of research on the impact of early grade success on future academic achievement, since most studies have attributed academic success to cultural identity and socio-economic factors, to the exclusion of the impact due to early academic performance,

(3) a lack of longitudinal cohort analysis of student performance under NCLB from Grade 3 to Grade 8 and Grade 10 respectively,

(4) a lack of multilevel research on the HSA data that takes into account both individual (i.e., student) and contextual (i.e., school) effects.

Early Academic Achievement

Much attention has been directed toward socio-economic status (SES) or minority status in explaining academic achievement. Such attention has intensified remarkably since the NCLB legislation was enacted. Under NCLB, the focus is on schools to close the academic achievement gaps between socio-economic groups, for example, between low SES students and high SES students, and between minority and White students. However, this may only account for a certain percentage of the variance in academic achievement. Most of the variance could instead be attributable to other individual-level characteristics, of which a hitherto largely overlooked factor is the students' early academic achievement. This early achievement is conceivably dependent upon the student's cultural identity and SES. To what extent early academic achievement may influence future academic performance over and beyond the effects of SES and culture is an important question largely neglected in the NCLB context. The present research was thus proposed based upon the premise that early academic achievement has its unique, and perhaps crucial impact on future academic achievement, an impact that has not been explicitly stressed in the language of NCLB but deserves serious attention from all stakeholders in public education.

Research Challenges

To establish the importance of early academic achievement on future academic performance, there is a need to include the usual socio-economic and cultural factors in the analysis at both the student and school levels. Moreover, there is a need to show not only that early academic achievement accounts for a considerable proportion of the variance in academic performance within elementary education, but also that its impact extends into middle and high school. The challenge is to model the effect with both socio-economic and cultural factors statistically controlled for at the student- and school-levels and to show that this impact remains significant from elementary to middle school and from elementary to high school.

In addressing the above challenges within the NCLB and Hawaii contexts, the current research investigated the impact of Native Hawaiian students' academic success at Grade 3 with White students as the control group. Three main research questions were addressed.

The first research question would address how Native Hawaiian students' early academic achievement at Grade 3 would impact their future academic achievement at

Grade 5 in comparison to White students with both socio-economic and cultural factors taken into account at the student- and school-level.

The second research question would seek to address how Native Hawaiian students' early academic achievement at Grade 3 would impact their future academic achievement at Grade 8 in comparison to White students, after controlling for socio-economic and cultural factors at the student- and school-level.

The third research question would seek to answer how Native Hawaiian students' early success at Grade 3 would impact their future academic achievement at Grade 10 in comparison to White students after socio-economic and cultural factors have been considered at the student- and school-level.

Due to the nesting of students within schools, a simple multiple regression model would not be appropriate to answer the first research question. Instead HLM would need to be adopted with both student- and school-level indicators, such as the student's free or reduced price lunch status, the percentage of free or reduced price lunch students in the school, the student's cultural background (whether Hawaiian or White) and the percentage of students with Native Hawaiian ancestry in the school.

The second research question presents an additional methodological challenge because cross-classification occurs when students from the same elementary school enter different middle schools. These students were cross-classified at two school levels, one elementary school and the other middle school. Hence, a cross-classified model would need to be used instead, with school-level predictors at the elementary level as well as at the middle school level. As the HLM used for answering the first research question, the student's free or reduced price lunch status, the percentage of free or reduced price lunch students in the school, the student's cultural background and the percentage of students with Native Hawaiian ancestry in the school would be included in the analysis.

Similarly, students in the same elementary school at Grade 3 were in different high schools when they reached Grade 10. Those students were also cross-classified at two school levels, one elementary school and the other high school. A regular HLM will not be appropriate to answer the third research question. Therefore, another cross-classified model would be required, with the same student- and school-level indicators included at the elementary and high school levels.

Multilevel analyses would be conducted for reading and math separately. Investigating the impact of early academic achievement within Hawaii's public schools would therefore require a careful synthesis of the HLM findings for different subjects and at different grade levels in search for a more or less consistent pattern across subjects and grade levels.

The above challenges may have so far hindered the research into investigating the impact of early academic achievement within the NCLB context. This research attempted to take on the challenges to develop and search for an underlying multilevel model to account for future academic achievement in reading and math within elementary schools, from elementary to middle school and from elementary to high school.

The No Child Left Behind Mandate

The predecessor of the NCLB mandate, the Elementary and Secondary Education Act (ESEA), was enacted in 1965, for the purpose of improving public education. The ESEA has served as the federal government's vehicle to address socio-economical disadvantages in public education, especially to help those who have been traditionally marginalized.

The ESEA was re-authorized in 2001 as NCLB. The reauthorization gives more power to the federal authority and provides stricter guidelines and sanctions. NCLB requires a single statewide accountability system for all public schools in each state. Under NCLB, reading and math performances on the state's assessments are important factors in determining whether a school has or has not met the NCLB yearly goals.

The NCLB mandate requires that each school report the percentages of students (a) participating in the NCLB assessment and (b) meeting proficiency in the following nine sub-categories for reading and mathematics (Zhang, 2009):

(1) All students

(2) Socio-economically disadvantaged students

(3) Disabled students

(4) Students with limited English proficiency

(5) Asian/Pacific Islander students

(6) Black students

(7) Hispanic students

(8) Native American students

(9) White students

The required rate of participation in statewide assessment is kept uniform at 95%. The proficiency rate is set by individual states, which is expected to increase yearly until it reaches 100% in 2014. Each school in Hawaii must meet the yearly target as shown in Table 1.1. Sub-categories of small sample sizes may be exempted; and for low performing schools, permission is occasionally granted for a lower yearly proficiency rate.

When a federally funded school continues to fail in meeting the state's established yearly proficiency objectives, sanctions will eventually be activated. The severity of the sanctions depends on the number of years a school has failed to meet the target, and can range from being labeled as needing improvement to the closing of the school (Zhang, 2009).

There have been considerable concerns, understandably, among public school staff directly affected by the federally imposed interventions and sanctions, especially those serving large proportions of disadvantaged and minority students. Through sanctions, NCLB allows the federal government to exert unprecedented influence on state, district, school and classroom practices.

The following listed actions are the consequences of not meeting the annual targets (Zhang, 2009):

(a) If the school misses the required targets for one or two years, it will be asked to improve from within and placed in the "school improvement"category.

(b) If the school misses the required targets for three or four consecutive years, it will be designated as "in need for improvement". The school will qualify for supplementary educational services (SES). Its students may be given a choice to transfer out of the school.

(c) If the school misses its targets for five consecutive years, the school will be subjected to "corrective action," which may range from receiving SES to decreased school decision-making regarding curriculum and pedagogy.

(d) If the school does not meet the required targets for six consecutive years, it will need to submit a plan for "restructuring."

(e) If the school misses the required yearly targets for the seventh consecutive year, it will need to implement its restructuring plan. Drastic actions may be taken as corrective measures, such as the dismissal of the principal, replacement of the staff, take-over by a private company, reorganization into a charter school or closing of the school.

NCLB represents a historical departure from previous educational legislations because it allows the federal government to enforce the rates of school improvement. Under NCLB, school characteristics are not considered when a decision is made to sanction. The NCLB mandate assumes that a school's progress can be measured by using the same yardstick (in this case, the annual targets) regardless of student- and school-level characteristics.

The focus of NCLB is to help disadvantaged groups by setting high academic standards, providing resources to underprivileged students through vouchers or supplemental educational services, and enforcing sanctions for failing schools. The federal government believes that through exercising these measures, it will close the achievement gaps between White and minority students, and between the non-disadvantaged and disadvantaged students.

Hawaii State Assessment

The Hawaii state legislature mandated, before NCLB, that the Hawaii Department of Education (HIDOE) should develop educational standards specific to Hawaii. The first edition of the Hawaii Content and Performance Standards (HCPS I) was produced in 1994, but in 1997-1998, in order to reduce the number of standards, HCPS I was revised. The second version of the standards was reorganized into six broad strands for reading and five for mathematics. The revised version has since been known as HCPS II. In 2006, HCPS II was revised to be HCPS III. Accordingly, the reading and math assessments were redesigned; and a science assessment was added.

Reading HSA covers three out of the six defined strands of content standards (Hawaii Department of Education, 2003b). These are:

(a) Comprehension processes (using strategies to construct meaning)

(b) Conventions and skills (applying linguistic and textual conventions for comprehension)

(c) Response (responding to a text from a personal, interpretive or critical stance)

Three other strands that are not assessed are:

(d) Range (various types of readings)

(e) Attitudes and engagement (confidence in and satisfaction with reading)

(f) Diversity (thoughtfulness about and respect for multi-cultural reading)

The Math HSA includes five strands (Hawaii Department of Education, 2003b):

(a) Numbers and operations (number system, computation and estimation)

(b) Measurement (attribute, unit, method and understanding of measurement)

(c) Geometry and spatial sense (dimensional property, visualization, coordinate geometry and transformation)

(d) Patterns, functions and algebra (numeric pattern, functional relationship and symbolic representation)

(e) Data analysis, statistics and probability (data organization, exploratory analysis, prediction and inference)

In 2002, students took the HSA at Grades 3, 5, 8 and 10. The raw scores for the reading or math assessments were scaled to a uniform range from 100 to 500 with 300 being the cut-off score for meeting proficiency. There are four categories of proficiency levels: (1) well below proficiency, (2) approaching proficiency, (3) meeting proficiency, and (4) exceeding proficiency. The scaled cut-off score of 300 is used consistently to rate students' performances across grades and across years. Table 1.1 shows HIDOE's annual achievement objectives for reading and math (Hawaii Department of Education, 2008).

Table 1.1 NCLB Annual Objectives for Hawaii

	2001 to 2004	2004 to 2007	2007 to 2010	2010 to 2012	2012 to 2013	2013 to 2014
Reading	30%	44%	58%	72%	86%	100%
Math	10%	28%	46%	64%	82%	100%

Seven years into the NCLB legislation, many of the schools are still finding it hard to meet the state's high expectations. Although criticisms and concerns have been voiced by administrators and teachers, HSA still stands out as being the first standards-based assessment system for Hawaii.

Despite the longitudinal perspective apparent in the annual objectives (see Table 1.1), there are two psychometric issues in assessing HSA performances over the years under NCLB. First, the items in HSA assessments are not vertically linked. A vertically linked set of items are items that can be placed on a common scoring scale that provides a measure for unidimensional content matter over different grade levels. Since HSA items are not vertically linked, it precludes the choice of growth modeling, whereby individual student performance slopes over time can be obtained to assess the increase in student knowledge in reading and math under NCLB. Second, HSA assessments were developed by different assessment contractors from 2002 to 2006 and from 2007 to 2009. This means that proficiency standards for the same subject matter at the same grade were set independently from the two periods 2002 to 2006 and 2007 to 2009. Since 2007 onwards HCPS III has been in effect. HSA scores cannot be compared before and after 2006. In other words, a standard cut-off score of 300 for proficiency may not be comparable from the years between 2002 to 2006 to the years between 2007 to 2009. This could also be a reason as to why no study has attempted to investigate the impact of early academic achievement in Hawaii because longitudinal analysis is not feasible given the data. Therefore, there seems to be a need to explore methodological approaches other than growth modeling if the long-term impact of early academic deficiency among students with Native Hawaiian ancestry is to be investigated.

It should also be pointed out that under NCLB Native Hawaiian students are grouped with students of Asian ancestry. Native Hawaiian students' achievement gap has been overlooked. Furthermore, although cultural identity and socio-economic factors have been prominently emphasized, insufficient attention has been directed to the impact of early grade performance on future academic achievement. This dissertation attempted to provide a method to develop multilevel models that focus on Native Hawaiian students' early academic achievement. To the best knowledge of the author this dissertation is the first study to provide a multilevel model to investigate how a minority group of students' early academic achievement affects their future academic performance under the NCLB legislation.

CHAPTER 2

REVIEW OF THE LITERATURE

The purpose of the literature search was to first discover any papers that focused on the impact of early grade success on future academic achievement in conjunction with the effects of socio-economic and cultural factors. A more focused search was subsequently conducted to discover empirical studies based upon large-scale assessments, similar in scope to the HSA. The search was limited to studies published in the period from 1994 to 2009 with a sample size of at least one thousand (except for one study, Chard et al., 2008). Relevant papers have been retrieved from the following databases: ERIC, Google Scholar, Questia, ProQuest and a number of journals. The search terms used are reading scores, math scores, large scale assessment, multilevel modeling, ethnicity, socio-economic disadvantage, Hawaii State Assessment (HSA), early grade success, early grade achievement, early childhood education, cohort analysis and at-risk groups.

Out of the 45 papers identified, 21 of the papers were not empirical studies and were excluded. 10 of the remaining 24 studies did not consider early grade as an explanatory factor in students' academic achievement. These ten studies were reviewed briefly in three sections, (i) culture's role in academic achievement (3 studies), (ii) poverty's effect on academic achievement (5 studies), and (iii) effects of culture and poverty (2 studies). The classification of the 10 studies into the three sections was based upon the authors' own perspectives of culture, ethnicity and poverty without considering the obvious association of socio-economic and cultural factors. The main focus of the review was on the other 14 empirical studies.

Disadvantages in Academic Achievement

Culture's Role in Academic Achievement

Culture has been traditionally defined as a set of behavioral characteristics that is common to a group of people who share similar traditions and history (Bodley, 1994). Race or ethnicity has often been used synonymously to represent culture. Culture and ethnicity are sometimes considered to be mutually inclusive, and the following review was based on the assumption that ethnicity provides a proxy for the influence of culture on academic achievement.

Academic achievement gaps have been studied extensively between various ethnic groups. A study in Texas (Hanushek & Rivkin, 2009) conducted descriptive and regression analysis of Black and White students' math scores. The study revealed that between third and eighth grade, there was a significant overall widening of the achievement gap between Black and White students in math. The students were then separated into two categories, those who had high performance at the third grade and those who had low performance at the third grade. The study showed that the achievement gap between Black and White students was more pronounced in the high performance group. The achievement gap between the two racial groups in the low performance category did not widen. This suggested there could be an interaction effect between socio-economic disadvantage and early success. However, no significance testing of the interaction effect was conducted. The authors reported that racial composition was found to account for much of the variability in the achievement gap. Furthermore, at the eighth grade, Black students in schools where the Black student population was less than 25% scored on average 0.15 standard deviations (SD) below the school mean. In schools where Black students constituted 25 - 50%, Black students scored on average 0.28 SD below the school mean, and in schools with a majority of Black students (above 50%), Black students scored on average 0.48 SD below the school mean. This differential effect due to minority status at the eighth grade, however, was not found in the third grade.

The above study did not investigate the impact of early success nor was a possible interaction effect with ethnicity tested. The nested structure of the data with students enrolled in schools was ignored in this study.

In Arizona, a statewide criterion-referenced test (CRT) known as the Arizona Instrument to Measure Standards (AIMS), has been administered in Grades 3, 5, 8 and 10 for reading and math since 2000. Using the ordinary least squares (OLS) regression analysis, Garcia (2007) calculated the annual percentage point change in proficient students by grade and subject from the state-level AIMS data. The study provided the percentage of students who either met or exceeded the AIMS proficiency level. Garcia found that the annual percentage point change in proficient students before and after the implementation of NCLB was higher for Native American students than other ethnic minorities in most grades and subjects since the implementation of NCLB. However, Native American students fell farther behind their White counterparts during the same period. In other words, even though both groups improved, the achievement gap actually widened after NCLB was implemented.

A recent report on Native Hawaiian students' performance from 1998 to 2000 in Hawaii's public schools was released in 2003 (Kana'iaupuni & Koren, 2003). According to this report, almost 79% of the schools where Native Hawaiian students accounted for more than 50% of the enrollment were in need of major improvement, as compared to only 17% of the schools in which Native Hawaiians accounted for less than 50% of the enrollment. Native Hawaiian students were ranked the lowest amongst all the major ethnic groups in Hawaii on both the reading and math Scholastic Aptitude Test (SAT) and SAT 9 (a shorter version of SAT, which was adopted from 1999). Only descriptive statistics were reported. There was no measurement of effect size for ethnicity. Native Hawaiian students' average scores were directly compared with the average scores of other ethnic groups.

The three studies were consistent in their findings regarding minority students' low academic performance. Since none of the studies took into account the poverty associated with ethnicity, these studies lack precision in identifying how much of the achievement gap was due to the cultural component after the economic factor had been partialed out.

Poverty's Effect on Academic Achievement

In the multilevel analyses reported by Hungi (2008), two separate three-level models were developed to account for achievement in reading and math respectively, based upon a sample of 72,376 students in Vietnam. In Hungi's models, pupils were at level 1, schools at level 2 and provinces at level 3. A composite variable of having one's own private corner in a home, learning materials at home, and parents' education was adopted as the measure of socio-economic status (SES). Among Grade 5 pupils, the standardized coefficients for the strongest predictors of math achievement were 0.16 for teachers' average score of content knowledge, 0.11 for SES at the individual level, 0.10 for SES at the school level, and 0.10 for average number of homework assignments corrected. The standardized coefficients for the strongest predictors of reading achievement were 0.12 for SES at the school level, 0.11 for SES at the individual level and 0.11 for the teacher's average score of content knowledge of reading comprehension. The effects of poverty at the individual level and school level were almost identical for reading and math. The impact of poverty seemed to be consistent for reading and math.

In their study conducted in the United Kingdom, Luyten, Peschar and Coe (2008) utilized multilevel models to investigate the Program for International Student Assessment (PISA) 2000 data in order to examine the grade effect on 6,327 students' reading scores. Two thirds of the students were in Grade 11 and the rest were in Grade 10. The grade effect was defined as the additional year of schooling from Grade 10 to Grade 11. The grade effect (one additional year from Grade 10 to Grade 11) was found to be a higher improvement in schools with a larger percentage of disadvantaged students. In other words, schooling had a greater positive impact on disadvantaged students' PISA reading scores. School-level SES accounted for most of the variance between the schools, and home language was found to be non-significant when gender and individual-level SES were controlled for in their analyses. SES in this study was derived by using a composite index consisting of parental occupation and parental education.

In his study of poverty's impact on academic achievement in Hawaii, Nochi (2008) used the HSA reading and math raw scores as the outcome variables in his multilevel analysis of Native Hawaiian students. He conducted eight separate analyses, four for reading and four for math at Grades 3, 5, 8 and 10, and found consistently that eligibility for free lunch decreased performances between 1.83 and 3.50 points in reading and between 1.60 and 4.24 points in math, on a raw score scale from 0 to approximately 70. An increase of 1% in the proportion of students eligible for free lunch at the school would decrease HSA scores between 0.06 and 0.10 points in reading and 0.08 and 0.13 points in math. Nochi did not find an overall significant effect of cross-level interaction between student- and school-level SES. In his study, student-level SES was found to be a strong predictor on academic achievement.

Pustjens, Van de gaer, Van Damme, Onghena and Van Landeghem (2007), in a study in the Netherlands, used a sample of 6,411 students in secondary schools and 5,927 students in primary schools in their three-level models, which had students at level 1, class at level 2 and school at level 3. Their study showed that student-level characteristics, especially SES, were able to explain most of the variance in academic achievement in Dutch or Math. Student-level variables explained around 70% of the variance in the academic achievement for students in the second year of secondary school. The effect size of student-level SES was 0.80 for math and 0.82 for Dutch. SES was a composite score combining family background characteristics such as parental education and income.

Van der Berg (2008) used two-level models to investigate what predictors affected the academic performance in reading and math of 3,163 South African students at the sixth grade. Principal component analysis of household and pupil possessions was conducted to develop an index for SES. Both school-level SES and student-level SES were found to be important predictors of student performance in reading or math. Moreover, interactions between individual-level SES and school-level SES were reported to be significant, and these two predictors interacted positively to produce higher achievements for students with high SES in high SES schools, i.e., students with higher SES will benefit more if they are placed in wealthier schools. In contrast, student-level SES had no impact within the poorer schools.

These five studies showed clearly that poverty had a negative effect on academic achievement. They were conducted in different contexts: one in Vietnam (Hungi, 2008), one in the United Kingdom (Luyten, Peschar, & Coe, 2008), one in the Netherlands (Pustjens, Van de gaer, Van Damme, Onghena, & Van Landeghem, 2007), one in Hawaii (Nochi, 2008) and one in South Africa (Van der Berg, 2008). These studies, however, did not consider the effect of ethnicity and thus ignored any cultural effect on academic achievement. One study (Pustjens et al., 2007) included home language in the analysis, which might be understood to have accounted for some cultural influence. Overall, those studies assumed that academic achievement disparity arises from poverty alone, without considering the association between poverty and ethnicity or culture.

Studies on the Effects of Culture and Poverty

Aikens and Barbarin (2008), in a study conducted in the United States, utilized a three-level growth model to develop trajectories of academic growth from kindergarten to Grade 3, based on the Early Childhood Longitudinal Study – Kindergarten (ECLS-K) data. A sample of 17,401students was included in the longitudinal kindergarten-to-third grade sample. In their study, school socio-economic disadvantage and school reading means contributed to student reading outcome more than family characteristics. Aikens and Barbarin (2008) used a composite score for student-level SES, which consisted of five components: the father and mother's education, occupation and household income. The White group was treated as the control group. In their study, ethnicity was found to have less impact on achievement than SES.

Driessen (2002) sampled 14,334 students in Grade 4 and 12,630 students in Grade 8 for his multilevel models to investigate students' verbal (Dutch) and mathematical achievements in Dutch schools. Parents' education was treated as a proxy for student-level SES. He found that parental education and ethnicity at the student level accounted for most of the variance in academic performance at the student level. At the school level, differences between school performances could mostly be attributed to the parental education and ethnicity. The effect of parental education was consistent regardless of the percentage of minority students in the school. In this study, the scale for SES may not be a precise measurement of poverty since the parents' education was used as a proxy for SES thus it may not entirely relate to poverty.

Aikens and Barbarin (2008) and Driessen (2002) showed that culture and SES may both be considered as significant factors that affect academic achievement. Although previous research (Abbot & Joireman, 2001; Harkreader & Weathersby, 1998; Saturnelli & Repa, 1995; Williams, 1972) has suggested poverty to be more influential than ethnicity in determining academic achievement, ethnicity and poverty are highly correlated. It is, therefore, important to consider the impact of poverty in conjunction with ethnicity (Patterson, Kupermidt, & Vaden, 1990; Peng & Wright, 1994; Wong & Alkins, 1999).

None of the ten studies reviewed above considered early grade success as an additional factor. The following section of the literature review covered 14 studies that included early grade success as a predictor.

Early Grade Success

Before the review of the 14 studies that included the early success predictor, a brief discussion of two papers (Kilian & Kagen, 1981; Milano, 1981) may be in order for theoretical reasons. Both papers focused on academic development between Grade 2 and Grade 5.

Milazzo (1981) maintains that there are critical points in a child's educational journey through the elementary grades that need to be addressed to improve the chances of student success in high schools. Finding at which grades these critical points must be addressed is important. With immediate improvements in these crucial grades, students may develop the foundational skills that will provide them with beneficial long-term effects. Milazzo has two premises: (a) educational problems that are identified in Grades 4 and 5 to middle school grades are the consequences of an accumulation of deficiencies at earlier grades, and (b) it is too soon to determine a lack of foundational skills before Grade 3. Thus, a logical place to identify academic deficiencies and implement appropriate changes in instructional practices to target underperforming students will be between the second and fourth grade. By diagnosing deficiencies between Grade 2 and Grade 4, educators may be able to come up with targeted instructional strategies to ensure future academic performance.

Kilian and Kagen (1981) studied the long-term effects of the Title 1 reading intervention. They found that students who were in the program in Grades 2 and 3 showed improvements in the third grade; however, the effect seemed to have diminished by the time they reached the sixth grade. Students who were identified as low achievers started to fall farther behind their peers after the third grade, and that the decrease coincided with the reduction in resources provided to the students over the years. Moreover, low-performing students who were not in the Title 1 reading intervention program also started to fall farther behind from Grades 3 and 4. It seemed important to provide low performing students with individualized help, especially during Grades 3 to 5. This, they claimed, would prevent low achieving students from severely falling further behind their peers in the later grades. Their findings have provided some empirical evidence for focusing on early academic achievement at Grade 3.

A detailed review of the remaining 14 studies is provided next. These 14 studies are listed in Table 2.1. One study (Chard et al., 2008) that had a sample size of 667 was included among the 14 studies due to its relevance to this dissertation even though its sample size was below the search criterion of 1, 000. The 14 studies were grouped into three categories, (i) impact of early grade as the only predictor, (ii) early grade's impact in conjunction with culture, and (ii) early grade's impact in conjunction with both cultural and socio-economic factors.

These studies used a range of analytical methods such as multilevel modeling, ordinary least squares (OLS) regression, logistic regression, multivariate methods, structural equation modeling (SEM) and descriptive analysis. These studies investigated the long-term effects of early grade success from kindergarten to elementary school, from lower to higher grades within elementary school, from elementary school to middle school, and from middle school to high school.

A summary of the 14 studies in chronological order starting from the most recent publication is provided in Table 2.1.

Impact of Early Grade as the Only Predictor

Liu and O'Connell (2005) utilized the ECLS-K data for the school year 1998-1999 for their early childhood study. Their final sample consisted of 3,534 kindergarten children. In their multilevel analyses, Liu and O'Connell found significant variation between children's initial kindergarten reading status as well as significant variation in children's reading growth patterns. Their multilevel analyses revealed that 86% of the estimated within-person variation could be attributed to the time effect. The variance of the true initial status and the growth variance among the students were found to be significant. There was a small positive correlation between initial status and change in reading ability over time. Thus, children with a higher initial reading status would have a slightly higher rate of growth in their reading achievement than students with a lower initial reading status. This finding suggests that early academic proficiency has a positive impact on future academic success as the learning rate for better students tends to increase more than students with lower early performance. However, poverty and cultural factors were not taken into account in their multilevel analyses. Although early academic achievement would impact later success, its unique effect could not be ascertained as the confounding socio-economic and cultural factors were not considered.

Table 2.1 Summary of 14 Studies

Author (Year)	Data	Student sample	Dependent variable	Independent variable	Significance	Method
Kieffer (2008)	ECLS-K longitudinal data	17,385 Kindergarten students from K1 to Grade 5	Reading score	EG: K SES: Provided by the United States Department of Education (NCES)	EG: Yes SES: Yes Ethnicity: No	Hierarchical linear modeling
Uyeno & Zhang (2007)	Hawaii, USA	8,940 students in reading, and 8,935 students in math	Proficiency status in reading and math	EG: Grade 5 SES: Eligibility for free or reduced price meal	EG: Yes SES:Yes Ethnicity: Yes	Logistic regression
Chatterji (2006)	ECLS-K longitudinal data	2,296 Kindergarten students from K1 to Grade 1	Reading score	EG: K SES: Provided by the United States Department of Education (NCES)	EG: Yes SES:Yes Ethnicity: Yes	Hierarchical linear modeling
Princiotta & Hausen (2006)	ECLS-K longitudinal data	9,796 Kindergarten students from K1 to Grade 5	Reading score and math score	EG: K SES: Provided by the United States Department of Education (NCES)	EG:Yes SES: Yes Ethnicity:Yes	Descriptive statistics
Chatterji (2005)	ECLS-K longitudinal data	2,296 Kindergarten students from K1 to Grade 1	Math score	EG: K SES: Provided by the United States Department of Education (NCES)	EG: Yes SES:Yes Ethnicity: Yes	Hierarchical linear modeling
Liu & O'Connell (2005)	ECLS-K longitudinal data	Cohort of 3,534 students from Kindergarten to first grade	Reading score	EG: K SES: Provided by the United States Department of Education (NCES)	EG: Yes	Hierarchical linear modeling
Takanishi (2005)	Hawaii, USA	11,773 students in Hawaii from Grade 3 to Grade 5	Proficiency status in reading and math	EG: Grade 3 SES: Eligibility for free or reduced price meal	EG: Yes SES:Yes Ethnicity: Yes	Hierarchical linear modeling
McNiece et al. (2004) British study	Cambridge exams longitudinal data	4,197 students in first sample and 3, 840 students in second sample	Reading and math scores	EG: End of primary school SES: Parents occupation	EG: Yes SES: Yes	Hierarchical linear modeling
Rathbun et al. (2004)	ECLS-K longitudinal data	10,500 students	Reading and math scores	EG: K SES: Provided by the United States Department of Education (NCES)	EG: Yes SES:Yes Ethnicity: Yes	Ordinary least squares regression
Burkam et al. (2004)	ECLS-K longitudinal data	3,664 students	Literacy, math and general knowledge	EG: K SES: Composite of parents' education, occupation and income	EG:Yes SES:Yes Ethnicity: Yes	Ordinary least squares regression, multivariate modeling
Chard et al. (2004)	Oregon and Texas, USA	667 students	Literacy skills	EG: K	EG:Yes Ethnicity: Yes	Structural equation modeling
Rugutt et al. (2002)	Louisiana, USA	11,627 students	Reading and math scores	EG: Grade 4 SES: Eligibility for free or reduced price meal	EG: Yes SES:Yes	Hierarchical linear modeling
Rugutt & Ellet (2001)	Louisiana, USA	26,051 students	Reading and math scores	EG: Grade 4 SES: Eligibility for free or reduced price meal	EG: Yes SES: Yes Ethnicity: Yes	Hierarchical linear modeling
Rugutt (2001)	Louisiana, USA	26,051 students	Reading and math scores	EG: Grade 4 SES: Eligibility for free or reduced price meal	EG: Yes Ethnicity: Yes	Hierarchical linear modeling

*EG refers to early grade predictor, SES socio-economic status predictor, and K Kindergarten.

Early Grade's Impact in Conjunction with Culture

In a longitudinal cohort analysis, Chard, Stoolmiller, Harn, et al. (2008) showed that demographic factors, such as ethnicity and home language, were found not to have a significant impact on the growth in oral reading proficiency (ORF) among 667 kindergarten students up to the third grade. First grade comprehension, ORF slope, academic competence, and Phonemic Segmentation Fluency (PSF) were found to be positive predictors of SAT-10 score, and behavioral problems and Black ethnicity were found to be negative predictors of SAT-10 at Grade 3. Students who were considered to be academically competent by their teachers but had low reading performance in the kindergarten or first grade did better than their counterparts who had low reading performance but were also considered to be academically less competent. The achievement gap between these two groups widened from kindergarten to third grade, indicating that early performance is important for later success. Their study, however, did not include school-level variables in the analyses nor did it control for the effect of SES on academic performance.

In Britain, McNiece, Bidgood and Soan (2004) utilized two multilevel models, one repeated measures (level-1was student scores, level-2 was students) and one three-level model (level-1 was students, level-2 was schools and level-3 was locales) to investigate how two generations of students performed from primary to secondary school. The first cohort consisted of students who were born during 1958 (*N* = 4,197), and the second sample consisted students who were born in 1970 (*N* = 3,840). Both cohorts had student records until the age of 16. The main finding from this study was that reading or math score at the end of primary school (ages 11 and 10) were found to be the strongest predictors of reading or math performance at the end of secondary school when they were at age 16. The results suggested that students who performed well at the end of primary school would perform well at the end of secondary school. This finding was consistent for both cohorts. Their multilevel analyses also showed that regional and local districts accounted for only 4% of the total variation in student performance at age 16. Upon closer examination, the individual level was actually a combination of student- and school-level information because the school level was omitted. Therefore any variance accounted for by the school-level variables might have been attributed to the student-level variables. Nonetheless, it was clear that early grade success at the end of primary school is an important predictor of student success in secondary school even after student ethnicity was taken into account. This study, however, did not consider either the student- or school-level SES effect, which would have separated the effect of early grade success from socio-economic disadvantage and thus improved the accuracy in estimating the impact of early success on later achievement above and beyond other socio-economic factors.

Early Grade's Effect in Conjunction with Culture and Poverty

Kieffer (2008) showed that initial English proficiency in kindergarten had a huge impact on fifth grade reading performance. By utilizing multilevel growth models on the ECLS-K data with a sample size of 17,385, Kieffer's research showed that language minority students with low initial English proficiency in kindergarten performed at a lower level in reading than Native English speaking students at the fifth grade. The growth curves of the language minority group and Native English speakers also differed significantly. However, language minority students who were comparable to Native English students in English proficiency in kindergarten had growth curves similar to those of Native English speaking students. This suggests that minority status does not impact future reading performance if the initial English language proficiency is on par with that of the Native English speaking students. With demographic factors such as SES (free or reduced priced lunch) and ethnicity, statistically controlled for initial English proficiency in kindergarten had a significant impact on fifth grade reading performance. Economic status alone could not adequately account for the impact of initial English proficiency on reading achievement at the fifth grade. This study also showed that for comparable socio-demographic factors, the growth trajectories tended to converge. Low English proficiency students scored 0.6 SD lower than Native English speaking students in kindergarten and 0.4 SD lower at the fifth grade. School poverty was found to moderate this impact. The difference in performances between the two groups was reduced in high poverty schools.

Uyeno and Zhang (2007) conducted the first cohort study in Hawaii on how early academic achievement in elementary school might influence future academic achievement in middle school under NCLB. Their logistic regression on a 2002 cohort of third graders investigated how HSA proficiency status at Grade 3 would affect proficiency status at Grade 7 for reading (N = 6,970) and math (N = 7,007). Their analysis was limited to four major ethnic groups in Hawaii's public education system: East Asian, Filipino, Native Hawaiian and White. Most of the students (77% for reading and 80% for math) did not change their proficiency status from Grade 3 to Grade 7 under NCLB. In their further analysis with the effect of ethnicity statistically controlled for, early academic achievement proved to have a unique and strong effect. Proficiency at Grade 3 would reduce the odds of failure at Grade 7 by at least 90% in either reading or math, after the students' cultural or socio-economic characteristics associated with ethnicity had been taken into account. However, their regression analysis did not accommodate cultural or socio-economic conditions at the school level. Uyeno and Zhang (2007) provided the clearest empirical evidence as of 2007 that early academic deficiency may need to be treated as a distinct and serious disadvantage in addition to the commonly acknowledged ethnic and economic disadvantages, if the NCLB objective is to be met at all.

Using the ECLS-K data comprised of a 2,296 kindergarten students nested in 184 schools Chatterji (2006) statistically controlled for child-level background differences such as SES and ethnicity, and found that the effect of kindergarten-entry reading level on first grade reading was 0.88, that is, for every 1 point increase in the kindergarten-entry score, there would be close to a 1 point increase in first grade reading outcome. Inclusion of kindergarten-entry performance improved the model's explanatory power from 5% to 38% as compared to when only socio-demographic factors were considered. Hence, early preparation prior to kindergarten seems critical for performing well at the first grade. At the school level, class size and teacher certification rate were significant correlates with reading at the first grade.

Princiotta and Hausken (2006) conducted a descriptive analysis of the ECLS-K data from kindergarten to fifth grade and found that White and Asian students scored higher on average than Black and Hispanic students in reading at the fifth grade. Kindergarten reading score was associated positively with academic achievement at the fifth grade. Approximately 65% of the students who scored in the highest third of the kindergarten cohort also scored in the highest third at the fifth grade, and 53% of the students who scored in the lowest third of the kindergarten cohort also scored in the lowest third at the fifth grade.

For math, Black and Hispanic fifth-graders performed lower on average than White or Asian students. Students who were below the poverty level also scored lower than students who were above the poverty level. Students' mathematics achievement score was found to be positively associated with kindergarten score. Of the students who scored in the highest third, 67% also maintained their position in the fifth grade and students who scored in the lowest third in kindergarten also scored lower in the fifth grade. Due to its descriptive nature, the study was limited to allowing only student-level variables to be considered thus excluding all contextual effects. Furthermore, this study examined the effects of initial academic status, poverty and ethnicity separately, thereby, not allowing the unique effect of each predictor to be identified. Although the overall results for initial academic status, poverty or ethnicity seemed to be strong, no significant testing was conducted. As confounding was not addressed in their study, their findings should be interpreted with caution.

Similar to her study on reading scores (Chatterji, 2006), Chatterji (2005) used the same data and same random-intercept multilevel models to investigate the math outcome in the first grade. The variance due to the school-level differences was around 20%, and about 35 to 40% of the variance in the first grade math achievement was attributed to student-level differences from beginning of kindergarten. For every 1 point increase in kindergarten-entry math, there would be close to a 1 point increase in first grade math.

All in all, student-level covariates accounted for around 43 to 64% of the total variance for predicting from kindergarten-end to first grade math. The early success predictor was found to account for the variance in later performance up to around 30 to 35%. As in her study on reading, she found kindergarten-level performance to be the most important predictor for performing well in first grade math above and beyond other socio-demographic predictors.

Takanishi (2005) used a two-level hierarchical linear model to explore 17 student-level and 6 school-level predictors expected to affect Grade 5 reading and math proficiency. The outcome of proficiency was set at proficiency in both reading and math. The cut-off score for proficiency was set at 300 on a scale from 100 to 500 for either reading or math. Although ethnicity and SES, at the student- and school-levels, were significant predictors, Grade 3 reading or math proficiency was found to be the strongest predictor. The odds of a proficient third grader meeting Grade 5 proficiency was 18.88 times of a non-proficient third grader. Takanishi's proficiency definition differed from the NCLB's guideline, which is specific to reading or math, but not to both subjects. It is not known how her results might be interpreted in regards to a student's NCLB status on reading and math separately.

Based on the ECLS-K to third grade data, Rathbun, West and Hausken (2004) showed that Black students had the lowest average mean achievement among the different ethnic groups for both reading and math in kindergarten and at the third grade. Third-grade Black students were also found to have an average reading gain score about 6 to 9 points and math gain scores about 9 to 14 points lower than their White, Hispanic, or Asian/Pacific Islander counterparts respectively. These findings were reported after variables such as gender, family characteristics, type of kindergarten program and school had been taken into account. Their findings showed that the achievement gap between Black students and the other ethnic groups had widened over the years. Hispanic students' raw scores were found to be significantly lower than their White and Asian counterparts; however, after controlling for early academic achievement, no significant difference was found. Therefore, low early academic achievement of Black and Hispanic students had a negative impact on their future performance even after socio-economic factors were taken into account. However, how much of the percentage of variance in Black and Hispanic students' future academic performances was accounted for by their early grade achievement above and beyond other predictors was not investigated.

Burkam, Ready, Lee, and LoGerfo (2004) utilized OLS regression and multivariate modeling on the ECLS-K data to investigate how social class impacted learning in literacy, mathematics and general knowledge during summer. Their research findings showed that after taking into account factors such as race, gender, home language and single-parent status, initial cognitive ability prior to summer was significant for increased summer learning. Children who repeated kindergarten also learned less over the summer. The gain in mathematics knowledge during summer was found to be strongly associated with cognitive status at kindergarten year. The findings showed that early success at the end of kindergarten after taking into account ethnicity and SES was crucial for academic success during summer. Their study, however, did not consider contextual effects. Hence, it is not known whether school-level variables influenced students to learn more during the summer. This study focused on the significance of the early success predictor, but no attention was directed toward investigating the unique proportion of variance accounted for by the early success predictor above and beyond the other predictors.

Rugutt, Ellett and Kennedy (2002) used multilevel models to study the growth patterns of Black students' math achievement on the Louisiana Educational Assessment Program (LEAP). Students were separated into two groups, those eligible for free or reduced price meals and those who were not. A total of 11,627 Black students who had data in Grades 4, 6 and 7 were included in the analysis. Descriptive analysis showed that students who were eligible for free or reduced price meals on average had lower math scores in all three grades than their counterparts who were not eligible for free or reduced price meals. The achievement gap widened between the two groups over time from Grades 4 to 7. On average, students who were eligible for free or reduced price meals started with 181.18 points and gained 10.30 points and students who were not eligible started, on average, with 186.89 points and gained 11.40 points. Hence, students who were not eligible started on average at a higher point, and their scores were even higher at later grades. Initial performance at Grade 4 had an impact on future performance at Grade 6 and Grade 7. Within the high SES group, the higher the initial performance, the higher the students performed at later grades. The achievement gap between the low performers and high performers was maintained throughout the grades, meaning the achievement gap remained the same from Grade 4 up to Grade 7. In other words, a student with a higher initial score would also have a higher score over the years than his or her counterpart who had a lower initial math score at Grade 4. Within the low SES group, Grade 4 performance was also found to have a significant impact on performance at Grade 6 and Grade 7 and the difference between high performers and low performers widened at

Grade 6 and further at Grade 7. The rate of learning for low SES high performers was greater than that of low SES low performers.

In the above study, the interaction effect between SES status and initial status performance at Grade 4 was not tested because separate multilevel analyses were conducted for the low SES group and the high SES group. A better approach would have been a single multilevel model involving the interaction. That would have allowed a significance test on the interaction between SES and early success. Nevertheless, within each group, the early grade predictor was found to be significant in explaining future performance.

In an earlier study, Rugutt and Ellett (2001) had used multilevel models to analyze the achievement of a cohort of 26,051 students from Grade 4 to Grade 7. Their results showed that the initial achievement gaps in language and math at Grade 4 between Black and White students remained stable up to Grade 7. Black students' math scores on the average increased 10.01 points per year while their language scores increased 12.42 points per year on the average. White students' math scores on the average increased 13.80 points per year and language scores increased 15.00 points per year. In this study, students' data in Grades 4, 6 and 7 was also used to construct growth projections for two comparisons. The first comparison was between Black and White students who were on free or reduced price meals, and the second comparison was between Black and White students who were not on free or reduced price meals. In either case, Black students' initial scores in Grade 4 were significantly lower than White students for both language and math. Performance on Grade 4 language and math was also found to have a significant impact on students' performance at Grade 6 and Grade 7 in language and math. Although early success was found to be significant in each of their separate analyses, a single multilevel analysis incorporating ethnicity and SES would have been a more efficient method to investigate future performance. This single analysis would have also allowed them to test several interaction effects, such as the interaction between ethnicity and initial status at Grade 4, interaction between ethnicity and SES, and interaction between SES and initial status at Grade 4. It would also have been possible to check if a single underlying multilevel model could be generalized across the two subjects, language and math.

Rugutt (2001) conducted a third study using the same dataset in which he used two separate multilevel analyses to investigate how Grade 4 performance influenced performances at Grades 6 and 7 for Black and White students. In both analyses, he found that for Black and White students, initial performance at Grade 4 had a significant impact at Grade 6 and Grade 7.

The above study also showed that the achievement gaps were stable over the time period from Grade 4 to Grade 7. Again, Rugutt could have analyzed these findings in a single multilevel analysis by having ethnicity as an additional predictor. It seems that the three studies involving Rugutt were done in stages, the first involving two separate analyses for Black and White students, then in another study, two separate analyses for Black students with low SES and Black students with high SES, and finally in another study, two separate analyses for Black and White students who had low SES and Black and White students who had high SES.

All in all, these studies showed very similar patterns of the impact of early grade performance on later grades. Instead of three separate studies, Rugutt et al. could have evaluated the importance of all three predictors, initial performance at Grade 4, ethnicity and SES in one single multilevel analysis. In this way, he could have found the percentage of variance accounted for by the early grade predictor after having the effects of ethnicity and SES partialled out, provided he had tested and found that there were no significant interaction effects between the early grade predictor and the other two predictors.

Summary of the Studies

Overall, the 14 studies have shown a significant impact of early academic achievement on future academic achievement, be it from kindergarten to the first grade, kindergarten to the fifth grade, within elementary schools, from elementary to middle school or from middle to the high school. These 14 studies are summarized in Table 2.2.

Table 2.2 shows the 14 studies with both student- and school-level variables and the length in prediction from early academic achievement to future academic achievement. The studies are organized in a chronological order and the table has nine columns. The first column has the authorship and year listed. The next two columns are to check whether the study included SES and culture in conjunction with the early grade success predictor in their analyses. The fourth column is to check whether hierarchical linear modeling (HLM) was employed, that is, whether their research designs took into consideration the nested structure of the data. The next five columns are to check the length of the cohort studies, in other words, to check whether the cohort analysis conducted was from kindergarten to elementary school, within elementary school, from elementary to middle school, from middle to high school or from elementary to high school. A cross (X) in each column indicates that the study included the attribute listed in the column heading. For example, if the study had included SES in its analysis, then an X is placed in the cell in the column of SES for that study.

Nine of the 14 studies included both individual and contextual factors. Four of the studies spanned from elementary to middle school, one cohort study from the end of primary school to secondary school and another study just within elementary school. The rest of the cohort studies were from kindergarten to first grade. Not a single study had a cohort analysis from elementary up to high school.

The current research addressed the lack of investigation for the impact of early academic achievement on future academic achievement within the NCLB context. It also addressed the lack of research on the academic achievement of students with Native Hawaiian ancestry under NCLB for the 2002 cohort up to 2009.

From a methodological perspective, the current study provided multilevel analyses for a cohort of students that had remained in the public education system in Hawaii. As the assessments were not vertically linked, this was an alternative to growth modeling in order to address the lack of longitudinal analysis on the topic.

Table 2.2 Attributes of 14 Studies

Authorship and Year	SES	Culture	HLM	Kindergarten to elementary school	Within elementary school	Elementary to middle school	Middle to high school	Elementary to high school
Kieffer (2008)	X	X	X	X (K to Grade 5)				
Uyeno & Zhang (2007)	X	X				X (Grade 3 to 7)		
Chatterji (2006)	X	X	X	X (K to Grade 1)				
Princiotta & Hausen (2006)	X	X		X (K to Grade 5)				
Chatterji (2005)	X	X	X	X (K to Grade 1)				
Liu & O'Connell (2008)			X	X (K to Grade 1)				
Takanishi(2005)	X	X	X		X (Grade 3 to 5)			
McNiece(2004) (British study)		X	X				X (Ages 11 to 16)	
Rathbun(2004)	X	X		X (K to Grade 3)				
Burkam (2004)	X	X		X (K to Grade 1)				
Chard(2004)		X		X (K to Grade 3)				
Rugutt(2002)	X	X	X			X (Grade 4 to 7)		
Rugutt & Ellet (2001)	X		X			X (Grade 4 to 7)		
Rugutt(2001)		X	X			X (Grade 4 to 7)		

X refers to presence of attribute listed in the column heading.

CHAPTER 3

RESEARCH QUESTIONS

Purpose of the Study

The study was proposed to investigate how early success at the third grade affects the later academic success of Native Hawaiian students. This study followed the Hawaiian and White student's academic performance through different grades as they changed school from elementary to middle school and to high school. As of now, there has not been a single study that has utilized multilevel models to track a cohort of students from an elementary grade up to high school. Even though the research focused on Native Hawaiian students, this study may well be the first multilevel cohort study to investigate the long-term effects of early academic success up to high school in conjunction with socio-economic and cultural factors and without the benefit of vertical linking.

Table 3.1 shows the cohort analysis periods and the student- and school-level covariates that will be included in the multilevel analyses for the current study.

Table 3.1 Attributes and Cohort Periods in the Current Study

Study with Early grade	SES	Culture	HLM	Within elementary school	Elementary to middle school	Elementary to high school
Current study	X	X	X	X (Grade 3 to Grade 5)	X (Grade 3 to Grade 8)	X (Grade 3 to Grade 10)

X refers to presence of the attribute in the column.

Research Questions

The first research question is:

(1) Within the span of elementary education, how does early success (in Grade 3) affect Native Hawaiian students' future academic achievement (in Grade 5) in comparison to White students, with individual and school factors taken into account in a multilevel model?

This question addresses how the impact of early grade on future academic achievement may vary from the White cohort to the Native Hawaiian cohort within elementary education. Interaction effects will be explored provided the basic model shows significance of the main predictors. No cohort studies have been found so far that provide significant cross-level and single-level interaction effects among early grade success, SES and culture.

The second research question is:

(2) From elementary to middle school, how does early success (in Grade 3) affect Native Hawaiian students' future academic achievement (in Grade 8) in comparison to White students, with individual and school factors taken into account in a multilevel model?

This question addresses how the impact of early grade success on future academic achievement may vary from the White cohort to the Native Hawaiian cohort from elementary to middle school. Although the research question seems similar to the first research question, it calls for a more challenging analysis as students are cross-classified into the two levels of schools. Interaction effects will again be explored. However, interaction effects may prove to be untenable due to sparseness of data, that is, too few students are classified into many, if not most, of the possible combinations of a particular elementary school and a particular middle or high school.

The third research question is:

(3) From elementary up to high school, how does early success (in Grade 3) affect Native Hawaiian students' future academic achievement (in Grade 10) in comparison to White students, with individual and school factors taken into account in a multilevel model?

This research question seeks to answer how the impact of early grade success on future academic success may change from the White cohort to the Native Hawaiian cohort from elementary to high school. This research question will require another cross-classified model as students are cross-classified at the two levels of schools.

Significance of the Study

Students in Hawaii are tested from Grade 3 to Grade 10. The 2002 third grader cohort is the first cohort to have completed the HSA assessments under NCLB. This study is therefore the first and complete cohort analysis. As a multilevel study, this research examined the performance of Native Hawaiian students in comparison to their White counterparts from the third through the tenth grade. This study may provide the most accurate estimate of the achievement gap between Native Hawaiian and White students, and determine how this gap varies through different stages of public education in Hawaii.

This study also investigated the possibility of Native Hawaiian students on par with White students at the third grade falling behind their counterparts at the later grades.

Additionally, this research may direct attention to the possibility that the low performance of Native Hawaiian students in the higher grades may not entirely be due to economic status and cultural factors. Native Hawaiian students' later academic problems might be to a larger extent due to their early deficiency with social and cultural factors statistically controlled for. Thus, it is important to investigate whether early success can account for a large proportion of the variance in their future academic performance that is not attributable to economic and cultural disadvantages. A careful investigation of how early HSA scores have affected future HSA scores among the first cohort under NCLB may guide early diagnosis of future academic deficiency and inform intervention at the individual level within the critical early stage of public education in addition to pedagogical adjustments based exclusively upon group identities.

The results of this research may be utilized to guide leadership in public educational administration to direct efforts to early instructional intervention for the Native Hawaiian student population, in order to improve their overall academic success up to graduation from high school.

CHAPTER 4

METHODS

Participants

The 2002 third-grade cohort of Native Hawaiian and White students were enrolled in 186 elementary schools from 15 complexes in seven districts. Table 4.1 shows the means and standard deviations for the Native Hawaiian and White students excluding special education (SPED) students. SPED students were excluded from the main analysis in this part of the study to prevent statistical confounding. English Language Learners (ELL) were also deleted from the cohort because there were only 23 Hawaiian ELL students and 15 White ELL students at the third grade. There were too few of them to support credible statistical modeling. However, SPED students were later included in the analyses reported in Chapter 6, in order to check the stability of the results with or without SPED students. The student population excluding ELL and SPED students hereafter is referred to as "regular" students.

Table 4.1 also presents, by ethnic group, the means and standard deviations for the HSA reading and math scores at Grades 3, 5, 8 and 10. Students who did not remain in the same school from Grade 3 to Grade 5 were excluded from the analysis for the HLM analysis.

For reading, starting with a sample of 4,757 students at Grade 3, the sample size was reduced to 3,361 at the fifth grade; at the eighth grade, the sample was further reduced to 3,077, and finally to 2,436 at the tenth grade. For math, the sample started from 4,757 and was reduced to 3,360 at the fifth grade; it was further reduced to 3,077 at the eighth grade and finally to 2,436 at the tenth grade.

Table 4.1 Descriptive Statistics for HSA Reading and Math for Regular Students

		Reading			Math		
Grade	Ethnicity	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
3	Hawaiian	2,994	274.73	56.23	2,994	237.59	54.76
	White	1,763	318.34	58.44	1,763	274.91	50.60
	All	13,562	285.75	62.97	13,563	247.75	60.61
5	Hawaiian	2,324	288.36	60.72	2,324	243.97	55.46
	White	1,037	329.15	61.05	1,036	278.41	57.67
	All	14,276	294.55	71.28	14,272	250.98	65.34
8	Hawaiian	2,333	300.89	31.85	2,333	270.12	30.86
	White	744	324.40	32.57	744	288.50	34.52
	All	13,015	306.68	36.96	13,016	275.99	37.55
10	Hawaiian	1,793	315.20	30.92	1,793	281.60	30.35
	White	643	336.15	27.70	643	301.91	34.15
	All	12,715	317.01	37.12	12,731	285.99	39.36

The descriptive statistics shown in the table are broken down according to the Hawaiian and White Cohort. The population descriptive statistics are provided under the heading All.

Although Table 4.1 shows that student performance has been steadily increasing, explaining these results in terms of an increase in performance for all students is misleading. There are two reasons for this. First, from 2002 to 2006, HSA was based on standards which were different from the modified standards for the years from 2007 to 2009. Thus, raw scores on the HSA reading or math assessments cannot be meaningfully compared across the years. Even though the statewide assessments in 2003 (Grade 3) and 2004 (Grade 5) were indeed based upon the same second revision of the Hawaiian Content and Performance Standards (HCPS II), the two assessments on the same subject were not vertically linked. In other words, each assessment in each year used a separate scale. Therefore, no cross-year comparison would be interpretable. The same is true of assessments from 2007 (Grade 8) to 2009 (Grade 10).

Table 4.2 School-Level Demographics

Grade	Year	ELL (%)	SPED (%)	SES (%)	Hawn (%)
3	2002	9	9	49	26
5	2004	5	12	48	26
8	2007	6	12	42	28
10	2009	7	12	39	26

Overall percentages for each grade within each category are shown for the respective years the HSA was administered.

Table 4.2 shows the percentages based on the proportions of students with ELL, SPED, low SES and Hawaiian ancestry, based on the average of the schools. A point to note is that these averages were not calculated directly from the individual-level data but obtained from the HIDOE website, i.e., http://arch.k12.hi.us/school/ssir/ssir.html. The two types of averages were different because calculating from the individual-level data would have only provided the averages of students who took the HSA assessments, whereas obtaining the averages from the HIDOE website provided averages in schools for the whole student population, i.e., all enrolled students who did or did not take the HSA assessments at the school.

Outcome and Predictor Variables

Reading or math HSA has two sections. Both sections contain multiple choice and constructed response items. HSA scores are based upon a scale from 100 to 500 with 300 as the cut-off for proficiency. The scale and cut-off scores are consistent for all grade levels and subjects. They were also consistent from HCPS II and HCPS III. However, it should be emphasized that no HSA scores allow for meaningful cross-year comparisons. Following is a list of the 17 variables that were used in this study:

Outcome Variables

There were six outcome variables, one for each multilevel model:

- HSA Grade 5 reading and math scores labeled as HSA 5_R and HSA 5_M respectively
- HSA Grade 8 reading and math scores labeled as HSA 8_R and HSA 8_M respectively
- HSA Grade 10 reading and math scores labeled as HSA 10_R and HSA 10_M respectively

Student-Level Predictors

Altogether there were five student-level predictors:

- HSA reading and math scores at Grade 3 labeled as HSA 3_R and HSA 3_M respectively and centered around the school mean
- Native Hawaiian ancestry labeled as Hawn and coded 1 if Native Hawaiian or 0 if White (This variable is based upon the designation at Grade 3.)
- Eligibility for the school lunch program labeled as SES and coded 1 if eligible or 0 if not eligible (Eligibility is based upon the designation upon Grade 3.)
- Special education labeled as SPED and coded 1 if SPED or 0 if not (SPED designation is also based upon Grade 3.)

In accordance with NCLB guidelines, socio-economic disadvantage was defined as family income at or below 180% of the state's poverty level. A student is allowed free lunch if his or her family's income is at or below 130% of the state's poverty level, and reduced price lunch if the family's income is at or below 180% of the state's poverty level.

School-Level Predictors

Six school-level predictors were included in the multilevel models. The school-level socio-economic predictor was the percentage of eligible students in the school, multiplied by 100. The school-level predictor of Native Hawaiian ancestry was represented as the percentage of students in the school, multiplied by 100. These percentages were centered around the mean of all schools.

- Percentage of low SES students in the elementary school (Grade 3) labeled as SchSES-3 and centered around the mean of all elementary schools

- Percentage of low SES students in the middle school (Grade 8) labeled as SchSES-8 and centered around the mean of all schools with middle grades
- Percentage of low SES students in the high school (Grade 10) labeled as SchSES-10 and centered around the mean of all high schools
- Percentage of Native Hawaiians in the elementary school (Grade 3) labeled as SchHawn-3 and centered around the mean of all elementary schools
- Percentage of Native Hawaiians in the middle school (Grade 8) labeled as SchHawn-8 and centered around the mean of all schools with middle grades
- Percentage of Native Hawaiians in the high school (Grade 10) labeled as SchHawn-10 and centered around the mean of all high schools

Overcoming Methodological Challenges

Three major methodological challenges had to be addressed in this study:

(1) There was no uniform scale linking the assessments from Grades 3 to 5, 5 to 8 and 8 to 10, thus rendering the conventional growth modeling untenable. This means that the growth projection could not be directly calculated. An alternative method to predict students' future success from early success was therefore required.

(2) Nesting of student data within elementary schools could be readily accommodated through a regular hierarchical model since most students in the data stayed in the same school from Grade 3 to Grade 5. However, nesting at the eighth or tenth grades presented a challenge as students in the same elementary school moved into different middle schools or high schools. Such data structure requires a cross-

classified model to predict the impact of early success at Grade 3 on achievement at Grades 8 or 10.

(3) Given the same predictors, many multilevel models were possible to predict the performance of reading or math at each of the Grades 5, 8 and 10. However, such an analytical approach would not have provided parsimonious or generalizable answers to the research questions. Thus, effort had to be made to search for a single multilevel model for investigating the impact of early success across the years for the 2002 cohort. Competing multilevel models had to be explored first to find the most consistent underlying pattern that could be used for predicting future academic achievement from early success. For that purpose the following three criteria: (a) criteria of feasibility, (b) generalizability, and (c) parsimony were adopted in the search for a generalizable multilevel model.

Model Specification

The criteria of (a) feasibility, (b) generalizability, and (c) parsimony in search of an overriding model are further explained next.

Feasibility

Feasibility is constrained by the amount and structure of available data. Since the current focus was on the long-term effects of academic success at Grade 3 on future academic performance, early grade success needed to be examined in conjunction with student and school characteristics. These confounding characteristics should be statistically controlled for. Developing a feasible model therefore required several considerations, the first of which was to consider confounding causal factors, such as ELL and SPED, before a legitimate interpretation could be obtained concerning the impact of early grade success. Thus, students who were ELL and SPED students were removed first from the data. Two analyses were then conducted, one without SPED students and one with SPED students included. Students who were ELL were also eliminated from the second analysis due their very small sample size.

The selection of predictor variables was based on the available data, such as characteristics at the individual and school levels provided by HIDOE. Only those predictor variables that were accessible from the HIDOE that were the closest proxies of the disadvantages were used in the analysis, the variable of the third grade reading or math HSA score as a proxy of early success, eligibility for free or reduced price meal as a proxy of socio-economic disadvantage, and ancestry as a proxy of culture. For school-level predictors, the percentage of students eligible for free or reduced price lunch was used as a proxy for school socio-economic disadvantage. The school-level percentage was not aggregated directly from individual students in the HSA data from HIDOE but was instead retrieved from a public HIDOE source on the World Wide Web at the following link http://arch.k12.hi.us/school/ssir/ssir.html, which provided more accurate school level information, because not all students sat for HSA. All individual-level binary variables such as Hawn, SES, SPED and ELL were initially included in exploratory multilevel analysis. School percentages of Hawn, SES, SPED and ELL were also included. However, ELL at the student- and school-level was later eliminated from the study due to the extremely low number of them among the Hawaiian and White students. The SPED variable was excluded from the analyses reported in Chapter 5, but was included in the analyses reported in Chapter 6. The two separate analyses were conducted to confirm that the multilevel models were able to provide stable results for when SPED was included or not. If both analyses showed similar results then a single underlying multilevel model could be used to investigate the impact of early success on future academic achievement for regular students or regular students with SPED students included in the population.

Furthermore, for Research Question One, only students who stayed in the same school from Grade 3 to Grade 5 were kept in the regular hierarchical modeling. For Research Questions Two and Three, cross-classified models were adopted to accommodate the cross-classification of students in two different levels of schools.

An initial exploration of the data structure for this research was conducted to investigate whether the "sparseness" of the data would allow random slopes or interactions in a crossed-classified model. Sparseness of data occurs when too few students are cross-classified, which results in many empty "cells", i.e., no students in some combinations of an elementary school and a middle or high school. A cross-classified model may not converge if too many such combinations contain no students. Although the data did show evidence of "sparseness" of data, slopes were still initially kept random across cells. Some of the models failed to converge or had errors such as negative variances. Thus, it was determined that keeping the slopes random across cells for all models was not feasible. A simpler cross-classified model with slopes kept fixed across the cells was adopted instead.

Generalizability

The second criterion for model specification is generalizability. Generalizability is constrained by the extent to which each assessment is specific to its own set of content and performance standards. Since there were three predictions from Grade 3 to Grade 5, from Grade 3 to Grade 8, and from Grade 3 to Grade 10 in two subjects, reading or math, there were altogether six analyses. The best fitting models in those analyses were likely to differ. Given the same predictors, any of the six analyses might be subject to the idiosyncrasies of the specific assessment. Therefore, the best fitting predictive model for one subject at one grade might be partly due to the overall pattern across subjects and grade levels and partly due to the idiosyncrasies of the particular assessment examined. This study was not intended to search for the best fitting model for one particular subject at a particular grade. Instead, this study was intended to search for an overriding pattern of results. If a generalizable model could be ascertained, then the consistent general model could then be employed to describe the effect of Native Hawaiian students' early academic achievement on their later success.

Parsimony

The third criterion of parsimony was followed to search for a model that would account for a considerable portion of the variance without being unnecessarily elaborate. Parsimony is dependent on by the number of same- or cross-level interactions allowed. Each hierarchical model has two options, either to treat the slopes as fixed, that is, not allowing the slopes to vary across schools; or random, that is, allowing slopes to vary across the schools. This choice is typically determined on the basis of whether or not a model can converge, whether variance estimates are significant or meaningful. One obvious problem is a model that converges to give a negative variance or a correlation above 1 or below -1. The eventual generalizable model must avoid such problems at each grade level and for each subject area. This necessarily requires searching for a simpler model.

Given the number of student-level predictors and number of school-level predictors, it would not be reasonable to start a complete model with all possible interactions. The exploratory analyses were first limited to main effects only and then expanded to two-way same-level or cross-level interactions to the exclusion of three-way or higher order interactions, which is a common practice in HLM. Any main or interaction effect that was found to be non-significant in most of the six analyses would be eliminated for the sake of parsimony. Retaining a consistently non-significant effect in a model would not have additional value for policy recommendations based upon the results.

To select the most parsimonious and most generalizable model, a series of analyses were initially conducted to determine if the slopes for each predictor across schools needed to be fixed. Next, under the same criterion, two-way same-level and cross-level interactions were examined to determine if there was a consistent pattern of interaction effects. If no consistent interaction effects were found, then interactions were eliminated from the models for the sake of parsimony. In other words, if it was determined that no same-level interaction or cross-level interaction was consistently significant across the models for reading or math, then the interaction effect was dropped in the final models.

After considering single-level and cross-level interactions, it was possible to find which model could best be generalized across the three predictive models from the third grade up to the tenth grade for reading or math. For example, if there was not even one interaction effect that was consistent across the three predictive models, then the conclusion would be that the main effects model was sufficient to explain a Native Hawaiian or White student's performance at Grade 5, 8 or 10 from Grade 3 HSA reading or math scores. Through this methodical approach to synthesizing the results, the most parsimonious model was sought to address the three research questions.

In Chapter 6, this parsimonious model was checked with the Native Hawaiian and White SPED third grade students included in the cohort. The results of the additional analysis, if consistent with the findings of the main analysis in Chapter 5, would prove that regardless of SPED status, the single underlying model would be capable of explaining the impact of early success on future academic achievement within the elementary, and up to the middle and high school years.

Based on the three criteria, (a) feasibility, (b) generalizability and (c) parsimony, all the resulting models were refined to arrive at a single underlying multilevel model to explain the effect of early success (in Grade 3) on future academic achievement (in Grade 5, Grade 8 or Grade 10).

Exploratory Analyses

Exploratory analyses were conducted in two stages, one stage to investigate main effects only with three steps as outlined in Table 4.3 and a second stage to investigate two-way interactions consisting of two further steps, one step for same-level interactions and the other step for cross-level interactions. The first stage of exploration was therefore conducted without any interaction effects in the multilevel models. Table 4.3 shows the variables that were included in the different analyses at this stage. In each of the analyses, all three student-level predictors were included. During the first stage there were three steps to explore. In the first step, only school-level SES was included in the analyses, and in the second step, only school-level percentage of Hawaiians and Part-Hawaiians were included in the analyses, and in the third step, both school-level SES and school-level percentage of Hawaiians and Part-Hawaiians were included in the analyses. The main effects models thus included 18 analyses, i.e., 2 subjects x 3 predictive models x 3 school-level combinations (SchSES by itself, SchHawn by itself, and SchSES and SchHawn together). The student-level predictors were kept in all of the 18 analyses, i.e., HSA 3, Hawn and SES.

Table 4.3 Predictors in the Main Effects Model for the Three Steps

Outcome variable	Student-level predictor included in all analyses	School-level predictor in first step	School-level predictor in second step	School-level predictors in third step
HSA 5	HSA 3 Hawn SES	SchSES-3	SchHawn-3	SchSES-3, SchHawn-3
HSA 8	HSA 3 Hawn SES	SchSES-3 SchSES-8	SchHawn-3 SchHawn-8	SchSES-3, SchHawn-3 SchSES-8, SchHawn-8
HSA 10	HSA 3 Hawn SES	SchSES-3 SchSES-10	SchHawn-3 SchHawn-10	SchSES-3, SchHawn-3 SchSES-10, SchHawn-10

The second column shows student-level predictors only, the next columns show the school-level predictors used in the three steps.

Analyses were also conducted to investigate whether it was feasible to allow the intercepts and slopes to randomly vary. Additionally, these analyses also allowed the investigation of which school-level variables were important to the model and which could be dropped without affecting the overall power of the model.

SchHawn-3, SchHawn-8 and SchHawn-10 could be dropped from the respective models without affecting the overall predictive power at the fifth, eighth or tenth grade. SchSES-8 and SchSES-10 were kept in the models due to the need of cross-classified models.

Following the initial main effects analysis, the intercepts and slopes were also checked for whether they could be kept random or had to be fixed. This test showed that the intercepts should be kept random but most of the slopes could be fixed across the predictive models. Furthermore, it was found that the variable SchSES-3 should be kept, as it was found to be significant in all the analyses.

The second stage of the exploratory analyses was conducted in two steps to focus on interactions. In the first step, all possible combinations of same-level two-way interactions at the student level were investigated. A total of 18 possible analyses (2 subjects x 3 interactions x 3 models) were conducted to see if there was a consistent pattern across the predictive models from Grade 3 to Grade 5, 8 or 10 across reading and math. Table 4.4 shows all possible combinations of the same-level two-way interactions that were investigated. The student-level predictors, i.e., HSA 3, Hawn and SES were kept in all of the analyses.

In the second step, all possible cross-level interactions were investigated. A total of 18 possible analyses (2 subjects x 3 interactions x 3 models) were conducted to see if there was a consistent pattern across the predictive models from Grade 3 to Grade 5, 8 or 10 across reading and math. Table 4.5 shows all possible cross-level two-way interactions between student-level predictors with SES at the school-level for reading and math at Grades 5, 8 and 10.

Table 4.4 Investigation of Interactions at the Student-Level

Outcome variables	Interaction at the student-level
HSA 5	HSA 3*SES, HSA 3*Hawn, SES*Hawn
HSA 8	HSA 3*SES, HSA 3*Hawn, SES*Hawn
HSA 10	HSA 3*SES, HSA 3*Hawn, SES*Hawn

Each same-level interaction was conducted sequentially for each grade level as shown. The first column corresponds to the outcome variable and the next column shows the student-level interactions that were tested for the three predictive models.

Table 4.5 Investigation of Cross-Level Interactions

Outcome variables	Interactions at the cross-level
HSA 5	HSA 3*SchSES-3, Hawn*SchSES-3, SES*SchSES-3
HSA 8	HSA 3*SchSES-3, Hawn*SchSES-3, SES*SchSES-3
HSA 10	HSA 3*SchSES-3, Hawn*SchSES-3, SES*SchSES-3

The second column shows each cross-level interaction tested for each of the three predictive models.

Those multilevel models were also investigated for whether random slopes could be set, several of the models failed to converge and hence the initial models were modified and tested for each random effect at a time, which showed that a fixed effects model for all the predictive models was acceptable. Results of these are further detailed in Chapter 5.

Analysis

There were two predictive models to answer each of the three research questions, one for reading and the other for math. To answer the first research question, a regular hierarchical model was developed to explain the variability in HSA Grade 5 reading or math by including HSA Grade 3 results, student-level SES, student-level ethnicity, school-level SES, and percentage of school-level Native Hawaiians as the predictors in the model.

The first of these two predictive models was intended to predict HSA reading success at Grade 5, and the second predictive model to predict HSA math success at Grade 5. These two analyses determined the effects of the predictors within elementary education, such as how early success in reading and math performance might affect the Native Hawaiian cohort's reading or math performance in comparison to the White cohort at Grade 5 after individual and contextual effects had been taken into account in a multilevel model.

A regular hierarchical model can thus be used to explain variability in Grade 5 reading or math. A basic hierarchical model with intercepts and slopes treated as random and no interaction effects is shown below.

Level 1 $Y_{ij} = \beta_{0j} + \beta_{1j}(\text{HSA } 3)_i + \beta_{2j}(\text{SES})_i + \beta_{3j}(\text{Hawn})_i + r_{ij}$

Level 2 $\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{Hawn-3})_j + \mu_{0j}$

$$\beta_{1j} = \gamma_{10} + \mu_{1j}$$

$$\beta_{2j} = \gamma_{20} + \mu_{2j}$$

$$\beta_{3j} = \gamma_{30} + \mu_{3j}$$

Reduced

$$Y_{ij} = \gamma_{00} + \gamma_{10}(\text{HSA 3})_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchHawn-3})_j + \mu_{0j} + \mu_{1j}(\text{HSA 3})_i + \mu_{2j}(\text{SES})_i + \mu_{3j}(\text{Hawn})_i + r_{ij}$$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
Y_{ij} = HSA Grade 5 score of the i^{th} student at the j^{th} elementary school

This model provided six fixed effects and five random effects. The six fixed effects were γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{01} and γ_{02} ; and the five random effects were μ_{0j}, μ_{1j}, μ_{2j}, μ_{3j} and r_{ij}. The restricted maximum likelihood (REML) estimation method was used for parameter estimation, and the Kenward-Rogers (KR) method was used to calculate the degrees of freedom (*df*) for the significance testing of the parameter estimates. An unstructured error matrix was used for the multilevel modeling.

Cross-Classified Model

To answer the second research question, a more complex analysis was required as there were two school-levels, the elementary school at the j^{th} level and the middle school at the k^{th} level. Some students remained in the same schools whereas others moved into different middle schools. Such data did not conform to the usual hierarchical data structure. Thus, unlike the analytic procedure for answering the first research question, a cross-classified hierarchical model would need to be adopted to explain the variability in Grade 8 reading or math.

There were two multilevel models to answer the second research question, one for reading and the other for math. The first model would predict HSA Grade 8 reading performance from Grade 3 reading performance, and the second predictive model would predict HSA Grade 8 math performance from Grade 3 math performance.

The following cross-classified model was tried out for the prediction from a student's Grade 3 performance to Grade 8 performance.

Level 1 $$Y_{i(jk)} = \beta_{0(jk)} + \beta_{1(jk)}(\text{HSA 3})_i + \beta_{2(jk)}(\text{SES})_i + \beta_{3(jk)}(\text{Hawn})_i + r_{i(jk)}$$

Level 2 $$\beta_{0(jk)} = \gamma_{00} + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-8})_k + \gamma_{03}(\text{SchHawn-3})_j + \gamma_{04}(\text{SchHawn-8})_k + \mu_{0j} + \mu_{0k}$$

$$\beta_{1(jk)} = \gamma_{10} + \mu_{1j} + \mu_{1k}$$

$$\beta_{2(jk)} = \gamma_{20} + \mu_{2j} + \mu_{2k}$$

$$\beta_{3(jk)} = \gamma_{30} + \mu_{3j} + \mu_{3k}$$

Reduced $$\begin{aligned} Y_{i(jk)} = {} & \gamma_{00} + \gamma_{10}(\text{HSA 3})_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i \\ & + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-8})_k \\ & + \gamma_{03}(\text{SchHawn-3})_j + \gamma_{04}(\text{SchHawn-8})_k \\ & + \mu_{0j} + \mu_{1j}(\text{HSA 3})_i + \mu_{2j}(\text{SES})_i \\ & + \mu_{3j}(\text{Hawn})_i + \mu_{0k} + \mu_{1k}(\text{HSA 3})_i + \mu_{2k}(\text{SES})_i \\ & + \mu_{3k}(\text{Hawn})_i + r_{i(jk)} \end{aligned}$$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
$k = k^{th}$ middle school
$Y_{i(jk)}$ = Grade 8 HSA score of the i^{th} student from the j^{th} elementary school and the k^{th} middle school

This model provided eight fixed effects and nine random effects. The eight fixed effects are γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{01}, γ_{02}, γ_{03} and $\gamma_{04;}$ and the nine random effects are μ_{0j}, μ_{1j}, μ_{2j}, μ_{3j}, μ_{0k}, μ_{1k}, μ_{2k}, μ_{3k} and r_{ij}. The REML estimation method was used for parameter estimation, and the degrees of freedom (*df*) for the significance testing of the parameter estimates was calculated using the KR method. The error matrix was unstructured.

To answer the third research question, a similar cross-classified hierarchical model was employed to explain the variability in Grade 10 HSA reading or math. There were two predictive models to answer the third research question. The first model would predict Grade 10 HSA reading performance from Grade 3 reading performance, and the second model would predict Grade 10 HSA math performance from Grade 3 math performance. This cross-classified model was very similar to the previous one, except for the outcome variable and one school-level predictor.

Level 1

$$Y_{i(jk)} = \beta_{0(jk)} + \beta_{1(jk)}(\text{HSA 3})_i + \beta_{2(jk)}(\text{SES})_i + \beta_{3(jk)}(\text{Hawn})_i + r_{i(jk)}$$

Level 2

$$\beta_{0(jk)} = \gamma_{00} + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-10})_k + \gamma_{03}(\text{SchHawn-3})_j + \gamma_{04}(\text{SchHawn-10})_k + \mu_{0j} + \mu_{0k}$$

$$\beta_{1(jk)} = \gamma_{10} + \mu_{1j} + \mu_{1k}$$

$$\beta_{2(jk)} = \gamma_{20} + \mu_{2j} + \mu_{2k}$$

$$\beta_{3(jk)} = \gamma_{30} + \mu_{3j} + \mu_{3k}$$

Reduced

$$\begin{aligned} Y_{i(jk)} = {} & \gamma_{00} + \gamma_{10}(\text{HSA 3})_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i \\ & + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-10})_k \\ & + \gamma_{03}(\text{SchHawn-3})_j + \gamma_{04}(\text{SchHawn-10})_k \\ & + \mu_{0j} + \mu_{1j}(\text{HSA 3})_i + \mu_{2j}(\text{SES})_i \\ & + \mu_{3j}(\text{Hawn})_i + \mu_{0k} + \mu_{1k}(\text{HSA 3})_i + \mu_{2k}(\text{SES})_i \\ & + \mu_{3k}(\text{Hawn})_i + r_{i(jk)} \end{aligned}$$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
$k = k^{th}$ high school
$Y_{i(jk)}$ = Grade 10 HSA score for the i^{th} student from j^{th} elementary school and the k^{th} high school.

This model provided eight fixed effects and nine random effects. The eight fixed effects were γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{01}, γ_{02}, γ_{03} and γ_{04}; and the nine random effects were μ_{0j}, μ_{1j}, μ_{2j}, μ_{3j}, μ_{0k}, μ_{1k}, μ_{2k}, μ_{3k} and r_{ij}. Parameters were estimated through the REML method, and the degrees of freedom (*df*) for the significance testing of the parameter estimates were again obtained through the KR method. Similarly, as in all of the above analyses, an unstructured error matrix was used.

The hierarchical linear modeling (HLM) was conducted using the MIXED procedure in the statistical software package SAS (Release 9.2).

CHAPTER 5

RESULTS FOR REGULAR STUDENTS

Intraclass Correlation Coefficients

The intraclass correlations (ICC), based on the unconditional means model, is usually the first step in HLM to determine to what extent the nested structure of the data accounts for the variation of the outcome variable. The ICC in this study was a measure of how much of the variance in the outcome variable was due to between-school differences. The ICC thus provided a rationale for considering a multilevel approach appropriate to separate the effects of student-level variables from those of school-level ones and produce more accurate parameter coefficients. Multilevel modeling is unnecessary if the ICC is close to zero. To calculate the ICC, an unconditional means model, also known as the intercepts only model, was used. In this most basic model, no specific predictors are included at the student- or school-levels. In the means model, the remaining variance is labeled as the residual, which is attributed to student-level differences only. In the example below, variance in the Grade 5 reading was partitioned into two components, at the school-level and at the student-level. The ICC for the fifth grade was calculated as follows:

$$\text{ICC} = \frac{\text{Estimated between-school variance at Grade 3}}{\text{(Estimated between-school variance + estimated between-individual variance)}}$$

$$ICC = \frac{386.27}{(386.27 + 3653.33)} = 0.096$$

The ICC of 0.096 means that 9.6% of the variance in Grade 5 reading scores could be accounted for by between-school differences, thus justifying a need for multilevel modeling for predicting Grade 5 reading outcomes. For the cross-classified model, the ICC was calculated in a slightly different manner. The ICC for eighth grade reading was calculated by partitioning the total variance into one component due to the between-school differences at the third grade and another component due to the between-school differences at the eighth grade. The separation of the total variance is demonstrated below:

$$ICC = \frac{\text{Estimated between-school variance at Grade 3}}{(\text{Estimated between-school variance at Grade 3 + estimated between-school variance at Grade 8 + estimated between-individual variance})}$$

$$ICC = \frac{29.97}{(29.97 + 73.92 + 1056.46)} = 0.027$$

This indicated that 2.7% of the variance in Grade 8 reading was due to between-school differences at the third grade.

$$ICC = \frac{\text{Estimated between-school variance at Grade 8}}{(\text{Estimated between-school variance at Grade 3 + estimated between-school variance at Grade 8 + estimated between-individual variance})}$$

$$\text{ICC} = \frac{73.29}{(29.97 + 73.92 + 1056.46)} = 0.073$$

This indicated that 7.3% of the variance in Grade 8 reading was due to between-school differences at the eighth grade.

The evidence also shows that between-individual variance outweighed between-school variance. A summary is provided in Table 5.1 for reading outcomes at the different grade levels.

Table 5.1 Variance Parameter Estimates and ICCs for Regular Students in Reading

Outcome	Grade Level variance	Estimate	*SE*	*Z*	*p*	ICC
Fifth	3rd Grade	386.27	65.56	5.89	<0.0001	0.096
	Residual	3653.33	91.39	39.98	<0.0001	
Eighth	3rd Grade	29.97	11.51	2.60	0.0092	0.027
	8th Grade	73.92	22.65	3.26	0.0011	0.073
	Residual	1056.46	27.72	38.11	<0.0001	
Tenth	3rd Grade	24.23	10.19	2.38	0.0174	0.024
	10th Grade	38.09	13.08	2.91	0.0036	0.040
	Residual	935.78	27.64	33.85	<0.0001	

The columns show the parameter estimates, the standard error (*SE*), the *Z* and the *p* values for the significance test. The last column shows the ICC at each grade level.

The ICC for math HSA scores at Grades 5, 8 and 10 were calculated in a similar manner for both the regular HLM and cross-classified models. Table 5.2 provides a summary of the variance parameter estimates in math. The variance parameter estimates in Tables 5.1 and 5.2 suggest that a predominant portion of the variance was attributable to the individual level. As shown in Table 5.3, only around 6.4% to 10.0% of the total variance in HSA reading or 6.1% to 12.4% of the total variance in HSA math, was attributable to between-school differences. Overall, school-level characteristics accounted for only about one tenth of the total variance in HSA performance in the multilevel models.

Table 5.2 Variance Parameter Estimates and ICCs for Regular Students in Math

Outcome	Grade level variance	Estimate	*SE*	*Z*	*p*	ICC
Fifth	3rd Grade	417.53	64.14	6.51	<0.0001	0.124
	Residual	2954.35	73.91	39.97	<0.0001	
Eighth	3rd Grade	22.10	10.14	2.18	0.0293	0.025
	8th Grade	78.13	23.08	3.39	0.0007	0.069
	Residual	993.08	26.17	37.95	<0.0001	
Tenth	3rd Grade	42.26	13.41	3.15	0.0016	0.036
	10th Grade	31.50	12.72	2.48	0.0133	0.025
	Residual	989.04	29.33	33.72	<0.0001	

The columns show the parameter estimates, the standard error (*SE*), the *Z* and the *p* values for the significance test. The last column shows the ICC at each grade level.

Table 5.3 Percentage of Variance due to School-Level Differences for Regular Students

Grade	Grade level variance	Reading	Math
5	3rd Grade	9.6%	12.4%
8	3rd Grade	2.7%	2.5%
	8th Grade	7.3%	6.9%
	Combined	10.0%	9.4%
10	3rd Grade	2.4%	3.6%
	10th Grade	4.0%	2.5%
	Combined	6.4%	6.1%

ICCs from Tables 5.1 and 5.2 for each grade level were multiplied by 100.

Results of Exploratory Analyses

The first stage of the exploratory analyses was done in three steps as explained in Chapter 4. The first step confirmed that the effect of school SES was significant in the models. The second step showed that the effect of school percentage of Hawaiians in the schools was found to be significant in some of the analyses when school SES was ignored. However, the third step revealed that when school SES was included with school percentage of Hawaiians, the effect of school SES was significant in all the models whereas the effect of school percentage of Hawaiians was no longer significant. A decision was made at that point to drop school percentage of Hawaiians as a school-level predictor for parsimonious modeling.

Results of the second stage of the exploratory analyses are described next. This stage involved two steps, one for analyzing student-level interactions and the next step for analyzing cross-level interactions. Table 5.4 shows the results for one of the student-level interaction effects across models and subjects as an example.

Table 5.4 Interaction of Hawaiian Ancestry with SES at the Student Level

	Hawn*SES	
Grade	Reading	Math
5	not significant	not significant
8	not significant	not significant
10	significant	not significant

Significance was based on $p < 0.05$. Each interaction effect for the three predicted models was tested.

For the only significant finding, the p-value was close to 0.05 ($p = 0.047$). Considering the large sample size, it should be easy for even a very small interaction effect to be shown as significant. The results thus suggest that relative to White students, the disadvantage of being a Hawaiian is generally consistent for low- or high-SES students. Separate analyses investigating the interaction effects of Hawaiian ancestry and early success (interaction effect of Hawn and HSA 3) showed non-significance in all of the multilevel models. Neither did the interaction effect between SES and early success (interaction effect of SES with HSA 3) turn out to be significant in any of the multilevel models.

Table 5.5 below shows an example of the investigation of a cross-level interaction, i.e., the interaction of Hawaiian ancestry (student-level) and school SES at Grade 3 (school-level). The cross-level interaction was non-significant across all multilevel models for either reading or math.

Table 5.5 Cross-Level Interaction between Hawaiian Ancestry and School-Level SES

	Hawn*SchSES-3	
Grade	Reading	Math
5	not significant	not significant
8	not significant	not significant
10	not significant	not significant

Significance was based on $p < 0.05$. Each interaction effect for the three predicted models was tested.

None of the cross-level interactions reached the 0.05 significance level in any of the multilevel models. Therefore cross-level interactions were excluded from further consideration. Since same- and cross-level interactions were found to be mostly non-significant, there seemed to be little statistical evidence for including any interaction effect in the final generalized multilevel models. In the final stage, a decision was made for the slopes to be kept fixed because the variation for the slopes across the schools was found to be non-significant in most of the cases. This led to the specification of the final generalized multilevel model with fixed slopes for the student-level predictors.

The final multilevel models are shown below for the three research questions respectively. For the first research question, a multilevel model including three student-level predictors and one school-level predictor is provided. In this model, all slopes were kept fixed, and no interaction effects were included.

Multilevel Model for Reading or Math at Grade 5

Level 1 $\quad Y_{ij} = \beta_{0j} + \beta_{1j}(\text{HSA 3})_i + \beta_{2j}(\text{SES})_i + \beta_{3j}(\text{Hawn})_i + r_{ij}$

Level 2 $\quad \beta_{0j} = \gamma_{00} + \gamma_{01}(\text{SchSES-3})_j + \mu_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20}$

$\beta_{3j} = \gamma_{30}$

Reduced $\quad Y_{ij} = \gamma_{00} + \gamma_{10}(\text{HSA 3})_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i + \gamma_{01}(\text{SchSES-3})_j + \mu_{0j} + r_{ij}$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
Y_{ij} = Grade 5 HSA score of the i^{th} student and the j^{th} elementary school

This model provided five fixed effects and two random effects. The five fixed effects were γ_{00}, γ_{10}, γ_{20}, γ_{30}, and γ_{01}; and the two random effects were μ_{0j} and r_{ij}. The restricted maximum likelihood (REML) estimation method was used for the parameter estimation, and the Kenward-Rogers (KR) method was used to calculate the degrees of freedom (*df*) for the significance testing of the parameter estimates. An unstructured error matrix was used for the multilevel modeling.

Cross-Classified Model for Reading or Math at Grade 8

Level 1 $\quad Y_{i(jk)} = \beta_{0(jk)} + \beta_{1(jk)}(\text{HSA 3})_i + \beta_{2(jk)}(\text{SES})_i + \beta_{3(jk)}(\text{Hawn})_i + r_{i(jk)}$

Level 2 $\quad \beta_{0(jk)} = \gamma_{00} + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-8})_k + \mu_{0j} + \mu_{0k}$

$\beta_{1(jk)} = \gamma_{10}$

$\beta_{2(jk)} = \gamma_{20}$

$\beta_{3(jk)} = \gamma_{30}$

Reduced $\quad Y_{i(jk)} = \gamma_{00} + \gamma_{10}(\text{HSA 3})_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-8})_k + \mu_{0j} + \mu_{0k} + r_{i(jk)}$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
$k = k^{th}$ middle school
$Y_{i(jk)}$ = Grade 8 HSA score of the i^{th} student from the j^{th} elementary school and the k^{th} middle school

This model provided six fixed effects and three random effects. The six fixed effects were γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{01}, and γ_{02}; and the three random effects were μ_{0j}, μ_{0k}, and r_{ij}. The REML estimation method was again used for the parameter estimation, and the KR method was again used to calculate the degrees of freedom (*df*) for the significance testing of the parameter estimates. An unstructured error matrix was also used in this analysis.

Cross-Classified Model for Reading or Math at Grade 10

Level 1	$Y_{i(jk)} = \beta_{0(jk)} + \beta_{1(jk)}(\text{HSA 3})_i + \beta_{2(jk)}(\text{SES})_i + \beta_{3(jk)}(\text{Hawn})_i + r_{i(jk)}$
Level 2	$\beta_{0(jk)} = \gamma_{00} + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-8})_k + \mu_{0j} + \mu_{0k}$
	$\beta_{1(j\ k)} = \gamma_{10}$
	$\beta_{2(j\ k)} = \gamma_{20}$
	$\beta_{3(j\ k)} = \gamma_{30}$
Reduced	$Y_{i(jk)} = \gamma_{00} + \gamma_{10}(\text{HSA 3})_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-10})_k + \mu_{0j} + \mu_{0k} + r_{i(jk)}$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
$k = k^{th}$ high school
$Y_{i(jk)}$ = Grade 10 HSA score of the i^{th} student from the j^{th} elementary school and the k^{th} high school

This model provided six fixed effects and three random effects. The six fixed effects were γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{01}, and γ_{02}; and the three random effects were μ_{0j}, μ_{0k}, and r_{ij}. Similarly, the REML estimation method was used for the parameter estimation, and the KR method was used to calculate the degrees of freedom (*df*) for the significance testing of the parameter estimates. The unstructured error matrix was adopted for this analysis. All the final multilevel models for this study were able to converge properly, that is, resulting in no negative variance or correlation with an absolute value larger than one.

HLM Results for Regular Students in Reading

Table 5.6 provides the descriptive statistics for each of the three analyses for the prediction for reading HSA from Grade 3 to Grade 5, Grade 3 to Grade 8 and Grade 3 to Grade 10 for regular students.

Table 5.6 Descriptive Statistics for Regular Students in Reading

		Grade 3			Grade 5/8/10		
	Ethnicity	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Grade 3 to Grade 5	Hawaiian	2,324	274.12	55.67	2,324	288.36	60.72
	White	1,037	318.03	58.95	1,037	329.15	61.05
Grade 3 to Grade 8	Hawaiian	2,333	269.72	53.32	2,333	300.89	31.85
	White	744	313.33	60.01	744	324..40	32.57
Grade 3 to Grade 10	Hawaiian	1,793	271.78	53.00	1,793	315.20	30.92
	White	643	314.21	60.44	643	336.15	27.70

The descriptive statistics for each of the three predictive models that had the complete data are shown in the table for reading.

To ensure that the loss of sample over the years did not impact the internal validity of the predictive analyses, the effect sizes of the difference between the Native Hawaiian and White cohort's third grade reading and math mean scores (Tables 5.6 and 5.8) were calculated. An example to show how the effect size was calculated for the prediction from Grade 3 to Grade 5 for reading is shown as follows:

Effect size = (Mean of White cohort – Mean of Native Hawaiian cohort) ÷
Standard deviation of the population at Grade 3(See Table 4.1)

= (318.03 – 274.12) ÷ 62.77 = 0.70

The effect sizes remained stable for predicting Grade 3 to Grade 5 (0.70 for reading and 0.61 for math), Grade 3 to Grade 8 (0.69 for reading and 0.61 for math) and Grade 3 to Grade 10 (0.67 for reading and 0.59 for math), showing that the achievement gaps at the third grade for the samples used in the three analyses did not differ for reading or math.

The parameter coefficients for the multilevel analyses for the three predictive models are presented in Table 5.7. Fixed and random effects are reported for HSA reading at Grades 5, 8 and 10. Standard errors are shown in parentheses. The R^2 which is the percentage of variance accounted for by all the predictors in the multilevel model is also provided. The R^2 served as a convenient measure of how accurately the multilevel model was able to explain the reading outcome at the fifth grade, eighth grade or tenth grade.

The intercept was the expected performance level of a White fifth grader who had his or her third grade reading score at the average of the schools, was not eligible for free or reduced price lunch at Grade 3, and was enrolled in a school with the average proportion of low SES students. This score was expected to be 306.27 points (95% confidence limit (CL): 302.98, 310.01) at the fifth grade. His or her score was expected to be 314.53 points (95 % CL: 311.90, 317.16) at the eighth grade. And his or her score was expected to be 328.01 points (95% CL: 325.38, 330.64) at the tenth grade. Such White students on average would exceed the cut-off score of 300.

An increase of 10 points at the third grade, with all other factors being constant, would translate to an increase of 8.00 points at the fifth grade (95% CL: 7.80, 8.20), an increase of 3.70 points at the eighth grade (95% CL: 3.50, 3.90), and an increase of 3.00 points at the tenth grade (95% CL: 2.80, 3.20).

A Native Hawaiian student would score lower than a White student by 6.16 points at the fifth grade (95% CL: 2.73, 9.59), 5.73 points at the eighth grade (95% CL: 3.42, 8.04) and 7.13 points at the tenth grade (95% CL: 4.60, 9.66), with all other factors being controlled for.

A student who was on free or reduced price lunch would have a score lowered by 4.67 points at the fifth grade (95% CL: 1.91, 7.43), 3.60 points at the eighth grade (95% CL: 1.64, 5.56), and 2.83 points at the tenth grade (95% CL: 0.58, 5.08), all other factors being constant in the model.

A student in a school which has 10% more students eligible for the school lunch program could be expected to score 7.30 points lower at the fifth grade (95% CL: 5.90, 8.70), 3.40 points lower at the eighth grade (95% CL: 2.80, 4.00) and 2.70 points lower at the tenth grade (95% CL: 2.10, 3.30), other factors being held constant.

R^2 or the percentage of variance accounted for by the model was calculated by correlating the predicted scores with the outcome scores. Although the R^2 decreased across the predictive models from 64% (Grade 3 to Grade 5) to 50% (Grade 3 to Grade 8) and to 38% (Grade 3 to Grade 10), the pattern was reasonable and the same effects remained significant across the models.

A warning may be in order here that direct comparisons across the grade levels would not be meaningful as the reading HSA between 2004 and 2006 were developed according to a set of standards different from that for the years from 2007 to 2009. The scores on the reading HSA, from 2004 to 2006, 2007 and 2009 were not based on the same scales.

Table 5.7 HLM Results for Regular Students in Reading

Year	2004	2007	2009
Grade	5	8	10
Fixed Effect	Coefficients	Coefficients	Coefficients
Intercept	306.27*** (1.91)	314.53*** (1.34)	328.01*** (1.34)
HSA 3	0.80*** (0.01)	0.37*** (0.01)	0.30*** (0.01)
Hawn	-6.16*** (1.75)	-5.73** (1.18)	-7.13*** (1.29)
SES	-4.67*** (1.41)	-3.60*** (1.00)	-2.83** (1.15)
SchSES-3	-0.73*** (0.07)	-0.34*** (0.03)	-0.27*** (0.03)
SchSES-8/10	N.A.	-0.10 n.s. (0.06)	-0.04 n.s. (0.05)
Random Effect	Variance Components	Variance Components	Variance Components
μ_{0j}	214.67*** (33.48)	14.75* (5.99)	3.73 n.s. (4.85)
μ_{0k}	N.A.	24.82** (8.66)	13.63* (6.27)
Residual	1521.57*** (34.83)	597.71*** (15.84)	634.73*** (18.67)
R^2	0.64	0.50	0.38

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.005$, n.s. non-significant,
N.A. not applicable

HLM Results for Regular Students in Math

The descriptive statistics for math are provided in Table 5.8 for each of the three analyses for the prediction in math HSA from Grade 3 to Grade 5, from Grade 3 to Grade 8, and from Grade 3 to Grade 10. HLM math results are presented in Table 5.9.

Table 5.8 Descriptive Statistics for Regular Students in Math

		Grade 3			Grade 5/8/10		
	Ethnicity	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Grade 3 to Grade 5	Hawaiian	2,324	237.63	53.57	2,324	243.97	55.46
	White	1,036	274.82	51.28	1,036	278.41	57.67
Grade 3 to Grade 8	Hawaiian	2, 333	233.09	52.12	2,333	270.12	30.86
	White	744	270.01	51.86	744	288.50	34.52
Grade 3 to Grade 10	Hawaiian	1,793	235.69	51.61	1,793	281.60	30.35
	White	643	271.65	51.25	643	301.91	34.15

The descriptive statistics for each of the three predictive models that had the complete data are shown in the table for math.

The intercept represented the expected performance of a White student who had a third grade math score at the average of his or her school, was not eligible for free or reduced price lunch at Grade 3 and was enrolled in a school with the average school performance and average proportion of free lunch students. This score was expected to be 259.07 points at the fifth grade (95% CL: 255.46, 262.28). The student would be expected to have a score of 280.36 points at the eighth grade (95% CL: 277.36, 283.36). At the tenth grade, this student would be expected to have a score of 293.86 points on the standardized HSA math assessment (95% CL: 291.12, 296.60). In all cases, such a White student was expected to fail to obtain the cut-off score of 300 and therefore would not be proficient in math at Grade 5, 8 or 10.

A student who scored 10 points higher than his or her peer at the third grade on the math HSA, with all other factors being constant, would score higher by 8.30 points at the fifth grade (95 % CL: 8.10, 8.50). Similarly, a student who scored 10 points above his or peer at the third grade would be expected to have a score of 4.00 points higher at the eighth grade (95% CL: 3.80, 4.20) and 3.50 points higher than his or her peer at the tenth grade (95% CL: 3.30, 3.70).

A Native Hawaiian student would score lower than a White student by 3.65 points at the fifth grade (95% CL: 0.65, 6.65), 4.78 points at the eighth grade (95% CL: 2.57, 6.99), and 6.83 points at the tenth grade (95% CL: 4.29, 9.47), other factors being equal.

A student who was eligible for free or reduced price lunch would have a decrease of 5.29 points at the fifth grade (95% CL: 2.70, 7.88), 3.92 points at the eighth grade (95% CL: 2.04, 5.80), and 2.92 points at the tenth grade (95% CL: 0.65, 5.19), with all other factors kept constant in the model.

A student in a school that had 10% more SES students at the third grade would be expected to have a score of 6.80 points lower at the fifth grade (95 % CL: 5.40, 8.20), 3.40 points lower at the eighth grade (95 % CL: 2.80, 4.00) and 3.60 points lower at the tenth grade (95% CL: 2.80, 4.40), with other factors being held constant.

As in reading, R^2 decreased over the years, i.e., 68% (Grade 3 to Grade 5), to 52% (Grade 3 to Grade 8) and 41% (Grade 3 to Grade 10). The pattern of significant effects was the same across the grade levels.

A similar warning is warranted that a direct comparison across grade levels would not be meaningful as the math HSA between 2004 and 2006 were also developed according to a set of standards different from that for the years from 2007 to 2009. Just as in reading, assessments from 2004 to 2006 and from 2007 to 2009 were not based on the same scales.

Table 5.9 HLM Results for Regular Students in Math

Year	2004	2007	2009
Grade	5	8	10
Fixed Effect	Coefficients	Coefficients	Coefficients
Intercept	259.07*** (1.84)	280.36*** (1.53)	293.86*** (1.40)
HSA 3	0.83*** (0.01)	0.40*** (0.01)	0.35*** (0.01)
Hawn	-3.65* (1.53)	-4.78*** (1.13)	-6.83*** (1.32)
SES	-5.29*** (1.32)	-3.92*** (0.96)	-2.92*** (1.16)
SchSESG3	-0.68*** (0.07)	-0.34*** (0.03)	-0.36*** (0.04)
SchSES-8/10	N.A.	-0.01 n.s. (0.07)	-0.02 n.s. (0.05)
Random Effect	Variance Components	Variance Components	Variance Components
μ_{0j}	278.40*** (38.77)	21.20*** (6.86)	19.54* (7.70)
μ_{0k}	N.A.	58.33*** (16.04)	13.76* (7.15)
Residual	1141.27*** (28.63)	542.57*** (14.48)	647.55*** (19.27)
R^2	0.68	0.52	0.41

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.005$, n.s. non-significant, N.A. not applicable.

Residual Reduction in HLM for Regular Students

The importance of Grade 3 performance was determined by calculating how much of the variance could be unambiguously attributed to it. This was accomplished through two residual reduction methods. The first is via residual reduction due to Grade 3 performance as shown in Table 5.10 and the second method is by computing the R^2 (the percentage of variance accounted for) which will be elaborated in the following section. Using the first method, the residual reduction attributed to the model without early success was found by subtracting its remaining error from the means model (with no predictors), the residual reduction attributed to the model with early success included was found by subtracting its remaining error from the means model. Residual reduction attributed to early success was then found by subtracting the residuals attributed to the model with and without early success as a predictor in the multilevel analysis as shown in Table 5.10.

Table 5.10 Residual Reduction for Regular Students

Outcome	Early success predictor	Residual variance reduction		
		Grade 5	Grade 8	Grade 10
Reading	HSA 3_R included	2303.36	523.07	337.01
	HSA 3_R not included	517.29	142.07	104.06
	Attributable to HSA 3_R	1786.07	381	232.95
Math	HSA 3_M included	1952.21	471.21	381.95
	HSA 3_M not included	389.75	98.06	106.77
	Attributable to HSA 3_M	1562.46	373.15	275.18

The residual reduction attributed to early success at each of the three grade levels for reading and math were computed by subtracting the model with the HSA 3 not included from the model with HSA 3 included.

An example is provided to show the calculation of the percentage of residual reduction in fifth grade reading due to early success above and beyond other predictors in the multilevel model. For Grade 5 reading, the percentage of residual reduction due to early success was computed as follows:

$$\frac{\text{Residual reduction due to early success}}{\text{Residuals in the models}} = \frac{\text{Residual accounted (HSA 3\_R included) - Residual accounted (HSA 3\_R not included, Table 5.10)}}{\text{Residual due to individual differences + Residual due to school level differences (Table 5.1)}}$$

$$= (2303.36 – 517.29)/(386.27 + 3653.33)$$

$$= 1786.06/4039.60$$

$$= 0.44$$

About 44% of the variance in fifth grade reading could be accounted for by HSA 3_R.

Similarly, the residual reduction due to early success was also computed for the cross-classified models. An example for calculating the error reduction for Grade 8 reading is shown below:

$$\frac{\text{Residual reduction due to early success}}{\text{Residuals in the models}} = \frac{\text{Residual accounted (HSA 3\_R included) - Residual accounted (HSA 3\_R not included, Table 5.10)}}{\text{Residual due to individual differences + Residual due to third grade + Residual due to eighth grade school differences (Table 5.1)}}$$

$$= (523.07 – 142.07)/(29.97 + 73.93 + 1056.46)$$

$$= 381/1160.36$$

$$= 0.33$$

About 33% of the variance in eighth grade reading could be accounted for by HSA 3_R. Table 5.11 summarizes the percentage of residual variance reduced by the early grade predictor at Grades 5, 8 and 10.

Table 5.11 Percentage of Residual Reduction for Regular Students

Grade	Reading	Math
5	44%	46%
8	33%	34%
10	23%	26%

The percentage for each residual reduction attributed to the early success predictor is shown for reading and math at each grade level.

The early success predictor was shown to be able to explain more of the variance in the outcome variable than any other predictor in all the models. The second method for determining the importance of early success was based upon R^2 familiar to most educators and researchers. R^2 was determined for when the HSA 3_R or HSA 3_M variable was included and when it was not included in the multilevel models while keeping all other predictors in the models. The increase in R^2 due to early success was then determined by subtracting the R^2 without early success from the R^2 with early success. The changes in R^2 due to early success were almost identical to the corresponding percentages of residual reduction reported in Table 5.11. Early success was found to be the most important explanatory predictor for long-term academic achievement. It should be noted that this unique effect due to Grade 3 performance was determined by two methods, analogous to the sequential partitioning of variance in non-hierarchical general linear modeling (GLM). A summary of the percentage of variance accounted for (R^2) by third grade alone is also provided in Table 5.13.

Table 5.12 Incremental R^2 for Regular Students

Subject	Early success predictor	Grade 5	Grade 8	Grade 10
Reading	HSA 3_R included	0.64	0.50	0.38
	HSA 3_R not included	0.19	0.18	0.14
	Attributable to HSA 3_R	0.45	0.32	0.24
Math	HSA 3_M included	0.68	0.52	0.41
	HSA 3_M not included	0.21	0.15	0.14
	Attributable to HSA 3_M	0.47	0.37	0.27

R^2 attributable to early success was obtained through subtracting the model without HSA from the model that included HSA.

Table 5.13 R^2 due to Early Success for Regular Students

Grade	Reading	Math
5	45%	47%
8	32%	37%
10	24%	27%

Percentages shown for each grade level was found by multiplying the R^2 attributable to early success from Table 5.12 by 100.

It was found that the early success predictor could account from about around 1.5 to 2.6 times the percentage of variance accounted for by other predictors in the models for HSA reading or math outcomes at the fifth, eighth and tenth grades. For example, the R^2 due to the unique effect of early grade for Grade 5 reading was 45% (Table 5.13). Without the early grade predictor, all the other predictors accounted for 19% (Table

5.12). Therefore, the unique variance attributable to early success at Grade 5 reading is around 45%/19% = 2.6 times that of the other predictors. This was clear evidence that compared to socio-economic or cultural factors early academic achievement is a predominant factor for future academic achievement.

The results of this study so far have shown that as far as regular education students are concerned, that early success is the strongest predictor of future academic achievement of Native Hawaiian and White non-SPED students. The HLM analyses revealed that the impact of school context has on student's academic achievement is lower than that of the characteristics of individual students. School-level differences accounted for about 6.1% to 12.4% in all the multilevel models. For each of the models examined, most of the variance could be attributed to individual-level differences. Individual predictors can explain more of the variability in student's future academic success than school-level predictors can. And among the student-level predictors investigated, early success has proved to be the most influential predictor of future academic achievement.

Standardized Coefficients for Regular Students

The next stage of the analysis was to examine the consistency of the effects of the student-level characteristics via standardized coefficients. Standardized coefficients (*Z*) are useful for interpreting the effects of the predictors across the grade levels. Raw coefficients can be converted to standardized coefficients (*Z*) by using the standard deviations of students' reading or math scores at the different grades. For example, at Grade 5, the standard deviation for reading was 71.28 (Table 4.1), the raw coefficient for

HSA 3_R of 0.80 (Table 5.6) was converted to a standardized coefficient of approximately 0.01, $Z = 0.80/71.28 = 0.01$. All raw coefficients were converted into standardized coefficients so that the stability of the effects might be checked across the grade levels. A Z of 0.1 for the coefficient of HSA 3_R would mean a student who scores 1 point higher in third grade reading, all other predictors being constant, would expect to see an increase of 1% of the standard deviation at Grade 5. More or less the same positive impact was evident at Grades 8 and 10. Through Z values, it was possible to check whether the effect would remain constant across the years, such as the student who scores 1 point above his or her peer in Grade 3 would consistently be about 0.01 standard deviations above his or her peer on the Grades 5, 8 and 10 reading or math HSA. Table 5.14 summarizes the raw and standardized coefficients for reading and math.

Table 5.14 Standardized Coefficients for Regular Students

Subject	Predictors	Grade 5		Grade 8		Grade10	
		Raw	Standardized (Z)	Raw	Standardized (Z)	Raw	Standardized (Z)
Reading	HSA 3_R	0.80	0.01	0.37	0.01	0.30	0.01
	Hawn	-6.16	-0.09	-5.73	-0.15	-7.13	-0.19
	SES	-4.67	-0.07	-3.60	-0.10	-2.83	-0.08
Math	HSA 3_M	0.83	0.01	0.40	0.01	0.35	0.01
	Hawn	-3.65	-0.06	-4.78	-0.13	-6.83	-0.17
	SES	-5.29	-0.08	-3.92	-0.10	-2.92	-0.07

The standardized coefficients (Z) for the student-level predictors were calculated for each of the three models. Both raw and Z are shown.

The standardized coefficients in Table 5.14 show that the effect of early success on the reading or math remained fairly stable at Grades 5, 8 and 10. The impact of socio-economic disadvantage was also fairly constant across the grades for both reading and math at Grades 5, 8 and 10. However, the impact of having Hawaiian ancestry seemed to have increased in both reading and math. The disadvantage of Hawaiian ancestry more than doubled from Grade 5 to Grade 10 for reading and almost tripled for math, with the *Z* decreasing from -0.09 to -0.19 in reading and from -0.06 to -0.17 in math. Native Hawaiian students have a distinct disadvantage that increases over their years in public education. This unique disadvantage for Native Hawaiian students has been identified separately from the associated disadvantages due to poverty.

CHAPTER 6

RESULTS FOR REGULAR AND SPED STUDENTS

In Chapter 5, the results were based on regular education students only. This chapter reports the results based upon the studies in regular and special education. The analyses in this chapter would provide the answer as to whether the results would be stable when SPED students were included. One more student-level predictor, SPED, was added to the multilevel models. The models in Chapter 6 were slightly different from those in Chapter 5.

Table 6.1 displays the descriptive statistics for the combined regular and SPED students. The descriptive statistics in Table 6.1 were based on the students who were enrolled in the different grades. Descriptive statistics for the students who had complete data for the cohort analysis are reported in Table 6.6.

Table 6.2 reports the parameter estimates for the unconditional means model and their corresponding ICCs for regular and SPED students in reading. Examples for calculating the ICCs have been provided in Chapter 5. With the additional predictor, SPED, the ICCs for the between-school differences were not too much affected. Thus the results shown in Table 6.2 are similar to those reported in Chapter 5.

The parameter estimates and ICCs for math outcomes for both regular and SPED students are reported in Table 6.3. These results are therefore almost identical to those reported in Chapter 5 (see Tables 5.1 and 5.2).

Table 6.1 Descriptive Statistics for Regular and SPED Students

		Reading			Math		
Grade	Ethnicity	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
3	Hawaiian	3,431	265.70	60.09	3,431	228.92	59.32
	White	1,980	310.64	63.49	1,980	269.07	54.84
	All	13,562	285.75	62.97	13,563	247.75	60.61
5	Hawaiian	2,645	279.04	65.88	2,646	236.06	59.56
	White	1,145	322.62	66.04	1,144	272.88	60.86
	All	14,276	294.55	71.28	14,272	250.98	65.34
8	Hawaiian	2,705	296.15	34.08	2,702	266.22	35.52
	White	834	320.62	35.47	831	285.24	36.26
	All	13,015	306.68	36.96	13,016	275.99	37.55
10	Hawaiian	2,091	310.46	33.95	2,096	277.88	31.79
	White	717	332.92	30.36	716	299.40	35.46
	All	12,715	317.01	37.12	12,731	285.99	39.36

The descriptive statistics shown in the above table are separated according to the Hawaiian and White Cohort. The population's descriptive statistics are provided under the heading All.

Based on the ICCs in Tables 6.2 and 6.3, most of the variance was attributable to the individual level. Only around 6.2% to 9.2% of the total variance in reading was attributable to between-school differences and around 6.9% to 11.0% of the total variance in math was attributable to the between-school differences. Those ICCs in Table 6.4 were very similar to the ones reported in Chapter 5 (see Table 5.3).

Table 6.2 Variance Parameter Estimates and ICCs for Regular and SPED Students in Reading

Outcome	Grade level variance	Estimate	*SE*	*Z*	*p*	ICC
Fifth	3rd Grade	435.40	73.40	5.93	<0.0001	0.092
	Residual	4312.76	101.42	42.52	<0.0001	
Eighth	3rd Grade	35.75	11.99	2.98	0.0029	0.026
	8th Grade	94.74	27.96	3.39	0.0007	0.064
	Residual	1174.46	28.73	40.87	<0.0001	
Tenth	3rd Grade	28.71	11.19	2.56	0.0103	0.024
	10th Grade	46.53	15.34	3.03	0.0024	0.038
	Residual	1107.32	30.41	36.41	<0.0001	

The columns show the parameter estimates, the standard error (*SE*), the *Z* and the *p* values for the significance test. The last column shows the ICC at each grade level for reading.

Table 6.3 Variance Parameter Estimates and ICCs for Regular and SPED Students in Math

Outcome	Grade level variance	Estimate	*SE*	*Z*	*p*	ICC
Fifth	3rd Grade	424.47	66.78	6.36	<0.0001	0.110
	Residual	3444.55	81.00	42.53	<0.0001	
Eighth	3rd Grade	30.24	11.16	2.71	0.0067	0.020
	8th Grade	82.98	25.19	3.29	0.0010	0.072
	Residual	1086.81	26.70	40.71	<0.0001	
Tenth	3rd Grade	41.27	13.09	3.15	0.0016	0.040
	10th Grade	29.98	12.60	2.38	0.0174	0.029
	Residual	1087.30	29.93	36.32	<0.0001	

The columns show the parameter estimates, the standard error (*SE*), the *Z* and the *p* values for the significance test. The last column shows the ICC at each grade level for math.

Table 6.4 Percentage of Variance due to School-Level Differences for Regular and SPED Students

Grade	Grade level Variance	Reading	Math
5	3rd Grade	9.2%	11.0%
8	3rd Grade	2.6%	2.0%
	8th Grade	6.4%	7.2%
	Combined	9.0%	9.2%
10	3rd Grade	2.4%	4.0%
	10th Grade	3.8%	2.9%
	Combined	6.2%	6.9%

ICCs from Tables 6.2 and 6.3 were multiplied by 100 for each grade level.

The following model was adopted to address Research Question One:

Level 1 $$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{HSA 3})_i + \beta_{2j}(\text{SES})_i + \beta_{3j}(\text{Hawn})_i + \beta_{4j}(\text{SPED})_i + r_{ij}$$

Level 2 $$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{SchSES-3})_j + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

Reduced $$Y_{ij} = \gamma_{00} + \gamma_{10}(\text{HSA 3})_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i + \gamma_{40}(\text{SPED})_i + \gamma_{01}(\text{SchSES-3})_j + \mu_{0j} + r_{ij}$$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
Y_{ij} = Grade 5 score of the i^{th} student at the j^{th} elementary school

Six fixed effects and two random effects were reported for this model. The six fixed effects were γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{40} and γ_{01}; and the two random effects were μ_{0j} and r_{ij}. In this model, an additional fixed effect was in place for the SPED predictor. The estimation method for the parameter estimates and the method for degrees of freedom (*df)* were REML and KR respectively. An unstructured error matrix was used in this analysis.

The cross-classified model below was specified for Research Question Two:

Level 1 $\quad Y_{i(jk)} = \beta_0 + \beta_{1(jk)}(\text{HSA } 3)_i + \beta_{2(jk)}(\text{SES})_i + \beta_{3(jk)}(\text{Hawn})_i + + \beta_{4j}(\text{SPED})_i + r_{i(jk)}$

Level 2 $\quad \beta_{0(jk)} = \gamma_{00} + + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-8})_k + \mu_{0j} + \mu_{0k}$

$\beta_{1(jk)} = \gamma_{10}$

$\beta_{2(jk)} = \gamma_{20}$

$\beta_{3(jk)} = \gamma_{30}$

$\beta_{4(jk)} = \gamma_{40}$

Reduced $\quad Y_{i(jk)} = \gamma_{00} + \gamma_{10}(\text{HSA } 3)_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i + \gamma_{40}(\text{SPED})_i + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-8})_k + \mu_{0j} + \mu_{0k} + r_{i(jk)}$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
$k = k^{th}$ middle school
$Y_{i(jk)}$ = Grade 8 HSA score of the i^{th} student from the j^{th} elementary school and the k^{th} middle school

In this cross-classified model, seven fixed effects and three random effects were estimated. The seven fixed effects were γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{40}, γ_{01} and γ_{02}; the three random effects are μ_{0j}, μ_{0k} and r_{ij}. This cross-classified model had an additional fixed effect to be estimated for the SPED. The REML method was employed for the parameter estimation and the KR method was used for determining the degrees of freedom (*df*) for significance testing. An unstructured error matrix was used in the cross-classified model as in the previous multilevel analyses.

A similar cross-classified model was adopted for Research Question Three:

Level 1 $Y_{i(jk)} = \beta_0 + \beta_{1(jk)}(\text{HSA } 3)_i + \beta_{2(jk)}(\text{SES})_i + \beta_{3(jk)}(\text{Hawn})_i + + \beta_{4j}(\text{SPED})_i + r_{i(jk)}$

Level 2 $\beta_{0(jk)} = \gamma_{00} + + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-10})_k + \mu_{0j} + \mu_{0k}$

$\beta_{1(jk)} = \gamma_{10}$

$\beta_{2(jk)} = \gamma_{20}$

$\beta_{3(jk)} = \gamma_{30}$

$\beta_{4(jk)} = \gamma_{40}$

Reduced $Y_{i(jk)} = \gamma_{00} + \gamma_{10}(\text{HSA } 3)_i + \gamma_{20}(\text{SES})_i + \gamma_{30}(\text{Hawn})_i + \gamma_{40}(\text{SPED})_i + \gamma_{01}(\text{SchSES-3})_j + \gamma_{02}(\text{SchSES-10})_k + \mu_{0j} + \mu_{0k} + r_{i(jk)}$

where $i = i^{th}$ student
$j = j^{th}$ elementary school
$k = k^{th}$ high school
$Y_{i(jk)}$ = Grade 10 HSA score of the i^{th} student from the j^{th} elementary school to the k^{th} high school

This cross-classified model for predicting from Grade 3 to Grade 10 provided seven fixed effects and three random effects. The seven fixed effects were γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{40}, γ_{01} and γ_{02}; and the three random effects were μ_{0j}, μ_{0k} and r_{ij}. The REML method was employed for the parameter estimation and the KR method was used for computing the degrees of freedom (*df*) for significance testing. An unstructured error matrix was also used in this analysis.

HLM Results for Regular and SPED Students in Reading

Table 6.5 provides the descriptive statistics for each of the three predictive analyses. These descriptive statistics differed from those reported in Table 6.1 because they consisted of the cohort of students who had complete data for the predictive models.

Table 6.5 Descriptive Statistics for Regular and SPED Students in Reading

		Grade 3			Grade 5/8/10		
	Ethnicity	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Grade 3 to Grade 5	Hawaiian	2,645	266.10	58.92	2,645	279.04	65.88
	White	1,145	311.10	63.80	1,145	322.62	66.08
Grade 3 to Grade 8	Hawaiian	2, 705	260.67	57.20	2,705	296.15	34.88
	White	834	305.28	65.34	834	320.62	35.47
Grade 3 to Grade 10	Hawaiian	2,091	262.08	57.49	2,091	310.46	33.95
	White	717	306.44	65.72	717	332.92	30.36

The descriptive statistics for each of the three predictive models that had the complete data are shown in the table for reading.

To verify that there was no systematic loss of sample, the effect sizes of the difference between the Native Hawaiian and White cohort's (including SPED students) third grade reading and math mean scores (Tables 6.5 and 6.7) were calculated as in Chapter 5. The effect sizes were stable for predicting Grade 3 to Grade 5 (0.71 for reading and 0.66 for math), Grade 3 to Grade 8 (0.71 for reading and 0.65 for math) and Grade 3 to Grade 10 (0.70 for reading and 0.64 for math), showing that the achievement gaps at the third grade for the samples used in the three analyses did not differ for reading or math. Parameter estimates are presented next in Table 6.6 for reading. Standard errors are shown in parentheses.

Table 6.6 HLM Results for Regular and SPED Students in Reading

Year	2004	2007	2009
Grade	5	8	10
Fixed Effect	Coefficients	Coefficients	Coefficients
Intercept	306.67*** (1.85)	314.78*** (1.33)	328.43*** (1.33)
HSA 3	0.80*** (0.01)	0.37*** (0.01)	0.30*** (0.01)
Hawn	-7.12*** (1.70)	-6.21*** (1.11)	-7.21*** (1.25)
SES	-4.30*** (1.49)	-3.49*** (0.95)	-3.38** (1.11)
SPED	-20.35*** (2.25)	-7.74*** (1.38)	-10.90*** (1.60)
SchSES-3	-0.74*** (0.06)	-0.35*** (0.03)	-0.28*** (0.03)
SchSES-8/10	N.A.	-0.10 n.s. (0.06)	-0.06 n.s. (0.05)
Random Effect	Variance Components	Variance Components	Variance Components
μ_{0j}	197.68*** (30.81)	17.69*** (5.99)	5.31 (4.64)
μ_{0k}	N.A.	27.11*** (9.40)	15.77* (6.51)
Residual	1612.80*** (37.92)	604.30*** (14.93)	664.67*** (18.34)
R^2	0.67	0.55	0.45

*$p \le 0.05$, **$p \le 0.01$, ***$p \le 0.005$, n.s. non-significant. N.A. not applicable

The intercept was the expected fifth grade performance of a non-SPED White student who was not eligible for free or reduced price lunch, had a third grade score at the average of his or her school and whose school had the average percentage of low SES students. This score was estimated to be 306.67 points at the fifth grade (95% CL:

303.04, 310.30). At the eighth grade, this student's score would be 314.78 points (95 CL: 312.17, 317.39). At the tenth grade his or her score would be 328.43 points (95% CL: 325.82, 331.04). Hence, the typical non-SPED White student in a school with the average percentage of low SES students would be expected to perform above the proficiency score of 300.

A 10-point advantage in reading at the third grade, all other factors being constant, would result in an advantage of 8.00 points higher at the fifth grade (95% CL: 7.80, 8.20). This student would also have an advantage of 3.70 points at the eighth grade (95% CL: 3.50, 3.90) and an advantage of 3.00 points at the tenth grade (95% CL: 2.80, 3.20).

A Native Hawaiian student would score lower than a White student by 7.12 points in reading at the fifth grade (95% CL: 3.79, 10.45), by 6.21 points at the eighth grade (95% CL: 4.30, 8.39) and by 7.21 points at the tenth grade (95% CL: 4.80, 9.66), all other factors being equal.

With all other factors kept constant, a low SES student at the fifth grade would score 4.30 points lower in reading than a high SES student (95% CL: 1.38, 7.22). This student would also score 3.49 points lower than the high SES student at the eighth grade (95% CL: 1.63, 5.35) and score 3.38 points lower at the tenth grade (95% CL: 1.20, 5.56).

A non-SPED student would score 20.35 points higher than a SPED student at Grade 5 (95% CL: 15.94, 24.76), 7.74 points higher at Grade 8 (95% CL: 5.04, 10.44) and 10.90 points higher at Grade 10 (95% CL: 7.76, 14.04), all other factors being equal.

A 10% increase in a school's poverty at the third grade would lower a student's score by 7.40 points at the fifth grade (95% CL: 6.22, 8.60), by 3.50 points at the eighth grade (95% CL: 2.90, 4.10) and by 2.80 points at the tenth grade (95% CL: 2.20, 3.40).

The results in Chapter 6 are almost identical to those reported in Chapter 5. The multilevel models developed in this research hence provide evidence of stability of parameter estimates for whether SPED students were included or not in the multilevel analysis for investigating the impact of early success.

HLM Results for Regular and SPED Students in Math

Results of the multilevel analysis for regular and SPED students in math are provided next. The descriptive statistics are shown in Table 6.7. The HLM results are provided in Table 6.8.

Table 6.7 Descriptive Statistics for Regular and SPED Students in Math

		Grade 3			Grade 5/8/10		
	Ethnicity	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Grade 3 to Grade 5	Hawaiian	2646	229.57	58.09	2646	236.06	59.56
	White	1145	269.29	55.09	1145	272.88	60.88
Grade 3 to Grade 8	Hawaiian	2702	224.27	56.80	2702	266.22	35.52
	White	831	263.65	56.80	831	285.24	36.26
Grade 3 to Grade 10	Hawaiian	2096	226.19	56.19	2096	277.88	31.79
	White	716	265.10	56.87	716	299.40	35.46

The descriptive statistics for each of the three predictive models that had the complete data are shown in the table for math.

Table 6.8 HLM Results for Regular and SPED Students in Math

Year	2004	2007	2009
Grade	5	8	10
Fixed Effect	Coefficients	Coefficients	Coefficients
Intercept	259.84*** (1.65)	280.08*** (1.52)	294.13*** (1.32)
HSA 3	0.82*** (0.01)	0.39*** (0.01)	0.35*** (0.01)
Hawn	-4.15* (1.47)	-4.76*** (1.08)	-7.18*** (1.25)
SES	-6.11*** (1.21)	-3.99*** (0.92)	-3.17*** (1.10)
SPED	-10.79*** (1.95)	-4.00*** (1.35)	-2.89 n.s. (1.60)
SchSES-3	-0.66*** (0.07)	-0.32*** (0.03)	-0.34*** (0.04)
SchSES-8/10	N.A.	-0.04 n.s. (0.07)	-0.01 n.s. (0.04)
Random Effect	Variance Components	Variance Components	Variance Components
μ_{0j}	268.30*** (36.39)	17.61*** (7.05)	20.90*** (7.40)
μ_{0k}	N.A.	59.92*** (16.72)	10.26 (5.53)
Residual	1187.40*** (27.95)	570.53*** (14.19)	656.33*** (18.15)
R^2	0.71	0.53	0.45

*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.005$, n.s. non-significant. N.A. – not applicable

The intercept was the expected performance of a non-SPED White student with high SES who had a Grade 3 math HSA score at the average of his or her school which had the average percentage of low SES students. The student would be expected to have a score of 259.84 points at Grade 5 (95% CL: 256.61, 263.07). He or she would score 280.08 points at Grade 8 (95% CL: 277.10, 283.06). And, this student would score 294.13 points at Grade 10 (95% CL: 291.54, 296.72).

Other factors being equal, a student scoring 10 points above his or her peer at the third grade would score 8.20 points above his or her peer at Grade 5 (95% CL: 8.00, 8.40). This same student would also be expected to score 3.90 points higher than his or her peer at Grade 8 (95% CL: 3.70, 4.10), and 3.50 points higher at Grade 10 (95% CL: 3.30, 3.70).

White students outperformed students with Native Hawaiian ancestry by 4.15 points on average at Grade 5 (95% CL: 1.27, 7.03). They also outperformed students with Native Hawaiian ancestry by 4.76 points at Grade 8 (95% CL: 2.64, 6.88), and outperformed students with Native Hawaiian ancestry by 7.18 points at Grade 10 (95% CL: 4.74, 9.63).

A low SES student would be expected to score 6.11 points lower than his or her counterpart with high SES at Grade 5 (95% CL: 3.74, 8.48). This student would also be expected to score 3.99 points lower at Grade 8 (95% CL: 2.19, 5.80) and 3.17 points lower at Grade 10 (95% CL: 1.00, 5.33).

With other factors kept constant, a SPED student would score lower than a non-SPED student by 10.79 points at Grade 5(95% CL: 6.97, 14.61), and lower by 4.00 points at Grade 8 (95% CL: 1.35, 6.65). However, there was no significant difference between SPED or non-SPED status in Grade 10 on the math HSA.

A student in a school with a 10% higher low SES students in the third grade would score lower by 6.60 points at Grade 5 (95% CL: 5.20, 8.00) than a student in a third grade school with the average percentage of low SES students. This student would also be expected to score 3.20 points lower at Grade 8 (95% CL: 2.60, 3.80) and 3.40 points lower at Grade 10 (95% CL: 2.60, 4.20).

A warning for interpreting the math HSA is also warranted because the HSA math assessments from 2004 to 2006 were developed based on different standards from those between the years from 2007 to 2009. Therefore, comparison between 2004 to 2006 and 2007 to 2009 would not be meaningful across the different years.

Just as the results for reading HSA for regular and SPED students were found to be very similar to those reported in Chapter 5, the results for the math HSA for regular and SPED students are also found to be very similar to the results reported in Chapter 5. Hence, the results show that the multilevel analyses are stable whether or not SPED students are included in the analysis.

Residual Reduction in HLM for Regular and SPED Students

Residual reduction was similarly calculated as in Chapter 5. Residual reduction refers to the amount of residual reduced due to the unique effect of early success predictor (Grade 3 reading or math). The residual reductions due to the early success predictor are reported in Tables 6.9. Although they are very similar to those found in Chapter 5, the residual reduction is slightly lower due to the additional predictor SPED in the multilevel models that was attributed the variance common to both early success and SPED.

The percentages of residual reduction due to the early success predictor are summarized in Table 6.10. Those percentages were found, however, similar in magnitude to the reported in Chapter 5 (see Table 5.10), were slightly lower because of the additional SPED predictor in the multilevel analyses.

Table 6.9 Residual Reduction for Regular and SPED Students

Subject	Early success predictor	Grade 5	Grade 8	Grade 10
Reading	HSA 3_R included	2937.68	655.85	496.81
	HSA 3_R not included	1137.39	281.13	241.50
	Attributable to HSA 3_R	1800.29	374.72	255.31
Math	HSA 3_M included	2413.32	551.97	470.86
	HSA 3_M not included	839.75	196.84	195.19
	Attributable to HSA 3_M	1573.57	355.13	275.67

The residual reduction attributed to early success at each of the three grade levels for reading and math are computed by subtracting the residual accounted for by the model without the HSA 3 from the model with HSA 3 included.

Table 6.10 Percentage of Residual Reduction for Regular and SPED Students

Grade	Reading	Math
5	38%	41%
8	29%	30%
10	22%	24%

The percentage for each residual reduction attributed to the early success predictor is shown for each of the grade levels for reading and math.

Following the residual reduction method, an alternative method was used to identify the R^2 due to the unique impact of early success. The resulting R^2 value were quite close to those found in Chapter 5 but slightly lower because of the additional predictor SPED in the multilevel models.

All in all, third grade performance alone accounted for an additional 20 to 40% of the variation in Grades 5, 8 and 10 above and beyond the effects of the socio-economic and cultural disadvantages, indicating that its unique effect was greater than that of the socio-economic and cultural variables. Third grade performance seemed to be the most important explanatory predictor of long-term academic achievement. The findings were therefore similar to the ones reported in Chapter 5. These findings revealed that the additional SPED predictor did not affect the results of the multilevel analyses.

Table 6.11 Incremental R^2 for Regular and SPED Students

Subject	Early success predictor	Grade 5	Grade 8	Grade 10
Reading	With HSA 3_R included	0.67	0.55	0.45
	HSA 3_R not included	0.28	0.26	0.24
	Attributable to HSA-3_R	0.39	0.29	0.21
Math	With HSA 3_M included	0.71	0.53	0.45
	HSA 3_M not included	0.29	0.22	0.20
	Attributable to HSA 3_M	0.42	0.31	0.25

R^2 attributable to early success was obtained through subtracting the model without HSA 3 from the model that included HSA 3.

Table 6.12 R^2 due to Early Success for Regular and SPED Students

Grade	Reading	Math
5	39%	42%
8	29%	31%
10	21%	25%

The above table shows the R^2 from Table 6.11 multiplied by 100

Standardized Coefficients for Regular and SPED Students

As in Chapter 5, standardized coefficients were calculated to allow a comparison of the effects across the different predictive models. Table 6.13 summarizes the standardized coefficients for regular and SPED students. Those standardized coefficients are very similar to the ones reported in Chapter 5.

Table 6.13 Standardized Coefficients for Regular and SPED Students

Subject	Predictors	Grade 5		Grade 8		Grade10	
		Raw	Standardized (*Z*)	Raw	Standardized (*Z*)	Raw	Standardized (*Z*)
Reading	HSA 3_R	0.80	0.01	0.37	0.01	0.30	0.01
	Hawn	-7.12	-0.10	-6.21	-0.17	-7.21	-0.19
	SES	-4.30	-0.06	-3.49	-0.09	-3.38	-0.09
	SPED	-20.35	-0.29	-7.74	-0.21	-10.90	-0.29
Math	HSA 3_M	0.82	0.01	0.39	0.01	0.35	0.01
	Hawn	-4.15	-0.06	-4.76	-0.13	-7.18	-0.18
	SES	-6.11	-0.09	-3.99	-0.11	-3.17	-0.08
	SPED	-10.79	-0.17	-4.00	-0.11	-2.89	-0.07

The standardized coefficients (*Z*) for the student-level predictors were calculated for each of the three models. Both raw and (*Z*) are shown.

As in Chapter 5, the values were consistent for the effects of SES and early success. This means that with regular and SPED students combined the effect of SES and early success remained fairly stable from Grade 5 up to Grade 10. However, as in Chapter 5, the negative effect of Hawaiian ancestry increased over time (from Grade 5 to

Grade 10), showing that the disadvantage more than doubles for reading and triples for math for the Native Hawaiian students in public schools. Again, this shows the importance of considering policy implications for the Native Hawaiian students due to their disadvantage increasing over time while they remain in public education. Since the findings are very similar to those reported in Chapter 5, it shows that the analyses conducted in the current study are generalizable with or without SPED students included in the multilevel analyses.

CHAPTER 7

DISCUSSION

Impact of Early Academic Achievement

Since NCLB's focus is on closing the achievement gaps between minority and non-minority students, between SPED and non-SPED students, between ELL and non-ELL, and between low SES and high SES students, interventions understandably have been oriented towards group identity. The findings in the current research show that the importance of the impact of early academic achievement on future academic performance may have been overlooked.

This research has found that individual performance at the third grade is critical for future academic success at the fifth, eighth and tenth grades. Two studies conducted in Hawaii prior to the current study (Takanshi, 2005; Uyeno & Zhang, 2007) have also shown a significant and dominant impact of early grade success on future academic achievement. Takanishi's multilevel analysis was, however, limited to elementary education (from Grade 3 to Grade 5). Uyeno and Zhang's logistic regression analysis on the other hand was limited to the span from elementary school to middle school (from Grade 3 to Grade 7) and ignored school-level factors. The current study is the first study that has clearly shown that the impact of early academic achievement at the third grade remains stable up to high school using a multilevel approach that took into account both student- and school-level factors. This effect has been quantified at the fifth, eighth and tenth grades respectively and has been found to be above and beyond other factors such as culture and poverty. The results show that the impact of early academic achievement on future academic achievement remains consistent for both the White cohort and the Native Hawaiian cohort after controlling for cultural and SES factors, with SES also considered at the school level.

Through the residual reduction method and R^2 method, the percentage of variance accounted for by early success was quantified at the fifth, eight and tenth grades, and the results show that the third grade reading or math performance accounted for more than two times the variance due to the other disadvantages, such as Native Hawaiian Ancestry, SES and SPED, combined in reading or math. Thus, poverty and cultural factors alone cannot adequately account for future academic achievement within elementary education, from elementary to middle grade education or from elementary to the high school education. Early academic achievement's impact clearly demands attention under the NCLB context. A student's early academic achievement needs to be considered carefully since its impact is above and beyond the effects of other factors such as culture and poverty.

Previous studies have also suggested the importance of early academic achievement up to the elementary or middle grades. Chatterji (2005, 2006), in her studies on reading and math performances, found that the impact of kindergarten achievement accounted for more variance in first grade reading or math scores than the socio-demographic factors, such as poverty and ethnicity combined. Kieffer (2008) found that language minority students who had low initial English proficiency in kindergarten also had lower reading performance than Native English speaking students at the fifth grade. Interestingly, language minority students who were comparable to Native English speaking students in

English proficiency at the kindergarten level had similar learning growth patterns to those of Native English speaking students and performed as well as Native English speaking students at the fifth grade. In his study, academic achievement of minority students at the kindergarten level was shown to be more important than minority status for achievement at the fifth grade after SES was controlled for. Together, Chatterji and Kieffer's studies show that the influence of early academic achievement was not only stable at the first grade but is also stable at the fifth grade.

Early academic achievement at Grade 4 has also been found to be a significant predictor for Grade 7 performance in reading or math for Black and White students, as well as for low and high SES students (Rugutt et al., 2002; Rugutt & Ellet, 2001; Rugutt, 2001). This suggests the importance of early academic achievement from the elementary grade up to the middle school. The current study, however, showed that early academic achievement at the third grade is stable not only within elementary education, but also from the third grade to the eighth grade, and subsequently from the third grade to the tenth grade.

A recent report based on descriptive analysis (Hernandez, 2011) showed that third grade reading performance impacted high school graduation rates. A third grader having just basic reading proficiency was four times less likely to graduate from high school as a more proficient student; a third grader who did not even have basic reading skills was six times as unlikely to graduate. Previous research has found that reading is related to phonological awareness and that these skills should be developed by the third grade (Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). Moreover, phonological processing has also been found to be associated with math computation skills (Hecht, Torgesen, Wagner, & Rashotte, 2001). Reading difficulties were found to be related to slower progress in math. At the third grade, it was found that these students were unable to compare the magnitudes of numbers, unable to identify numbers proficiently and did not have sophisticated counting strategies (Gersten, Jordan, & Flojo, 2005). Emphasizing phonological awareness and focusing on reading difficulties associated with math deficits at the third grade could be pivotal for helping low performing third graders for future academic achievement. Thus, focusing on language development skills at the early grades may also be more beneficial for the Native Hawaiian students.

The purpose of this research, however, is not to downplay the importance of ethnic or socio-economic disadvantages but to show that an important factor such as early success should not be neglected if the NCLB goal is to help all students achieve long-term academic success. To attain the overall goal of NCLB, early academic achievement needs to be recognized as a crucial factor. More attention needs to be directed toward students' foundational academic preparation.

Disadvantage of Native Hawaiian Ancestry

Research conducted prior to the current study has shown that students with Native Hawaiian ancestry were the lowest performing group on standardized tests in Hawaii (Kanaʻiaupuni & Koren, 2003). This makes them the most at-risk group in Hawaii. However, their analyses were based on descriptive statistics; factors such as poverty and early academic achievement were not examined. In the current research, the impact of having Native Hawaiian ancestry has been quantified from Grade 3 to Grade 5, Grade 3 to Grade 8 and Grade 3 to Grade 10. The effect of having Native Hawaiian ancestry in this study showed that for reading or math, the disadvantage for a student with Native Hawaiian ancestry doubles as they progress from Grade 3 to Grade 10 for reading and triples for math in Hawaii's public schools, after early academic achievement at the third grade and SES had been controlled for. In other words, a Native Hawaiian student who was performing as well as his White counterpart at the third grade for reading or math, had the same SES status and who was enrolled in the same school would fall behind at the fifth grade, farther behind at the eighth grade, and even farther behind at the tenth grade.

Two studies support the current study's findings. Garcia (2007) found that Native American students fell farther behind their White counterparts during the NCLB period although both groups showed improvement in comparison to their performance before NCLB. A study in Texas (Hanushek & Rivkin, 2009), through descriptive and regression analysis of Black and White students' math scores, also revealed that between the third and eighth grades, the difference between these two groups of students widened. The two studies suggest that the achievement gaps between minority and White students became larger. These studies, however, did not control for early academic achievement in their analyses. The current study has shown that even after controlling for early grade achievement at the third grade and SES at the student- and school-levels, the negative impact of Native Hawaiian ancestry increases from elementary to high school in Hawaii's public schools. To my knowledge, there has not been a single study that has shown the increasing disadvantage of culture from the early grades up to high school after early academic achievement and poverty have been statistically controlled for.

With these findings in mind, teachers should pay particular attention to students with Native Hawaiian ancestry, even for those Hawaiian students who seem to be performing as well as the White students at the third grade, since these students would also fall behind their White counterparts at the later grades.

Impact of Poverty

The effect of SES for both regular and SPED students remained stable from third grade to tenth grade after controlling for early grade achievement at the third grade and Native Hawaiian Ancestry. The standardized coefficients for SES at the student- and school-levels were also found to be very similar to those reported in a study in Vietnam (Hungi, 2008). However, unlike Hungi's study, the current study's findings on poverty's effect were found after controlling for both early academic achievement and cultural factor. Nochi (2008), too, found that student-level SES was a strong predictor of academic achievement for students with Native Hawaiian ancestry and was stable at the third, fifth, eighth and tenth grades in his multilevel analyses. Nochi did not find significant cross-level interaction between student- and school-level SES similar to the finding in the current study that also showed no significant cross-level interaction in the multilevel models. The current study, however, utilized a cohort from the third to the tenth grade and showed the stability of the negative impact of poverty from the third grade to the tenth grade after controlling for early grade performance and culture. Nochi's study on the other hand focused on a single year across different grades and thus did not control for early success. The current study has thus confirmed that the impact of poverty remains stable from the third to the tenth grade in Hawaii's public schools.

Poverty also seems to impact academic achievement across different nations and cultures. Studies in the Netherlands (Pustjens et al., 2007), South Africa (Van der Berg, 2008), United Kingdom (Luyten et al., 2008), Vietnam (Hungi, 2008) and Hawaii (Nochi, 2008) have all shown the significant impact of poverty on student academic performances. Previous studies such as Aikens and Barbarin (2008) and Driessen (2002), which controlled for cultural factors, also found the impact of poverty to be significant. These studies therefore agree with the empirical evidence in the current study. However, none of the previous studies have shown the persistence and stability of the poverty effect for a cohort from elementary grades up to high school as reported in the current study with early success and culture statistically controlled for.

Generalizability

School-level differences reported in Chapter 5 for regular students and in Chapter 6 for regular and SPED students were almost identical. The impact of early success remained stable from third grade up to the tenth grade, whether SPED students were included or not in the analysis. The residual reduction due to early success was also similar, except that the residual reduction was slightly larger in Chapter 5 than those found in Chapter 6 because an additional predictor, SPED, was included in the multilevel analysis in Chapter 6. Any shared variance between the third grade predictor and SPED was therefore attributed to SPED causing a slight reduction in the variance attributed to the third grade performance. The second method to calculate the residual reduction was R^2, i.e., the percentage of variance accounted for by the third grade predictor also produced similar results for when SPED students were included or not in the analysis.

The values were almost identical to the ones reported in Chapter 5. The R^2 in Chapter 5, however, was slightly larger than the R^2 in Chapter 6 due to the additional SPED predictor included in the multilevel analysis since any overlap between SPED and early success was attributed to SPED. Thus, the results in the current study have shown that the impact of early academic achievement not only remains stable from third grade up to the tenth grade but it is also a dominant factor in predicting future success. Furthermore, these results were consistent for reading or math whether SPED students were included or not in the analysis. Thus, a single underlying multilevel model could explain the future performances of students at the fifth, eighth and tenth grades from their third grade performance.

School Differences under NCLB

The current study showed that only around 10% of the total variation in student academic performances at Grade 5, Grade 8 and Grade 10 in reading and math is attributable to school differences. Chatterji (2005, 2006) also reported that at the first grade, between-school differences accounted for only about 20% of the variance in student performances for reading or math. Since the between-school variance found in the current study is only 10%, it raises questions about whether allocating a large part of the responsibility to schools for solving students' low academic performance is fair. A larger percentage of the variation in academic performances were actually due to student-level characteristics, thus focus may need to be directed at the student level instead. NCLB's focus on school performances must be considered carefully in light of the findings in the current research. Between-school differences also seemed to decrease at the middle grades and further decreased at high schools, from 9.6% at Grade 5 to 7.3% at Grade 8 and to 4.0% at Grade 10 for reading; and 12.4% at Grade 5 to 6.9% at Grade 8 and to 2.5% at Grade 10 for math. When SPED students were included in the analysis, a similar pattern was obtained, the between-school differences decreased from 9.2% at Grade 5 to 6.4% at Grade 8 and to 3.8% at Grade 10 for reading; and from 11.0% at Grade 5 to 7.2% at Grade 8 and to 2.9% at Grade 10 for math. However, this finding must be considered cautiously because the reductions in the between-school differences at the later grades were found after the between-school differences at Grade 3 was included in the analysis.

To illustrate the importance of carefully considering school differences under NCLB, we can take the following example of comparing two schools from two different locales, such as a school in Hawaii Kai and a school in Waianae. The school in Hawaii Kai may have performed better than the school in Waianae, judging by school means. This comparison would lead us to believe that the school in Waianae is underperforming. However, two schools would also have very different student populations. The school in Waianae would likely to have more low SES students, more students with low academic achievement at the early grades, and more students having Native Hawaiian ancestry than the school in Hawaii Kai. Since the majority of the variance in academic performance is at the student level, then sanctioning a school seemingly low-performing based on the mean performance without considering individual variance is akin to saying a school with a higher percentage of high SES students, higher percentage of high performing early graders, and fewer students with Native Hawaiian ancestry deserves better recognition.

As the current research has shown that student-level differences may account for most of the variation in future academic achievement, then not considering early grade performance levels of the third graders seem to judge schools unfairly at the later grades. Under NCLB early grade achievement levels are not considered when schools are held accountable at later grades. Under NCLB guidelines, this would mean that low performing schools are required to improve their student performances at a faster rate than other schools.

This oversight has created a perceived disparity in achievement among schools under NCLB, since students with low early academic achievement at early grades and schools with the higher numbers of low performing early graders, higher proportion of low SES students and more students of Native Hawaiian Ancestry are subject to severe sanctions simply because they work with the most disadvantaged subpopulations.

The findings in the current research therefore suggest that because school-level differences only account for around 10% of the variance in student performances, focus needs to be directed at the individual-level characteristics more than school-level characteristics. The intent of this research is not to state that schools do not make any difference to student performance. The intent is, however, to increase the awareness that schools have limitations that could very well be beyond their control such as having students who may have not been prepared well at the early grades, having more low SES students and more students with Native Hawaiian ancestry.

Limitations

In this study, the partitioning of variance at the class level was not possible because information for classrooms was not available from HIDOE. This could have explained how much of the outcome variance at the fifth, eighth and tenth grades for reading or math was due to class-level differences. Since this variance could not be modeled in the current study, it was included in the residual variance instead.

Teacher characteristics were also not provided by the HIDOE. This information could have helped to answer whether teacher characteristics are able to explain how some students can improve faster than other students. Pedagogical methods employed by these teachers would have further provided insights toward addressing the needs of students with weaker foundational academic skills. Other socio-demographics like parent's education, income and family resources were also not available from HIDOE. Thus, controlling for such socio-demographics was not possible when accessing the impact of early success on future achievement for the 2002 cohort.

The cumulative effect from earlier grades such as the cumulative impact from the third grade combined with the impact from the eighth grade on the tenth grade reading or math performance could not be modeled in the current study. Investigating the cumulative impact would have created a methodological challenge which was beyond the ability of the software to analyze. Such an investigation would have required an analysis of a data structure which was cross-classified at two levels as some students moving into the eighth grade may not remain in the same school as in Grade 3, these students would be cross-classified at two school levels, one at the elementary and one at the middle school. Further, if some of these students moved into another school at Grade 10, it would have added another level of cross-classification. This then creates two levels of cross-classification instead of one. For the cumulative impact to be investigated, the cross-classified data with two cross-classifications had to be taken into account. Because modeling of such a data structure was not possible, a cross-classified data analysis with only one level of cross-classification was modeled instead. This precluded the investigation of the cumulative impact of earlier successes on future academic achievement.

Conclusion

In conclusion, this research has addressed the impact of early success on the Native Hawaiian students by considering them as a distinct group under NCLB. This research has also provided a method to analyze longitudinally the 2002 cohort of students' reading or math performance from third grade up to the tenth grade spanning a period from 2002 to 2009. The current research has also provided a single underlying multilevel model that took into account both individual and contextual effects in addressing the impact of early academic achievement on future academic success for Native Hawaiian students in comparison to their White counterparts.

The current study is therefore the first cohort study to have investigated the impact of early academic achievement at the third grade on future academic achievement at the fifth, eighth and tenth grades in a multilevel analysis controlling for culture at the student-level, and poverty at the student-level and school-level. The current study has also quantified the disadvantage of having Native Hawaiian ancestry under the NCLB period for the 2002 cohort. It is the first study to show that cultural factors such as Native Hawaiian ancestry confers a disadvantage on students that doubles for reading and triples for math from the third grade up to the tenth grade after controlling for early academic achievement at the third grade and poverty levels at the student- and school-levels. This research, therefore, has provided evidence that early academic deficits need to be considered as a separate disadvantage above and beyond other known disadvantages such as culture and economic status. The current study also showed that the variance due to between-school differences is smaller than the variance due to student-level differences thus informing that early academic performances of students need to be taken into account when judging schools at the later grades.

Although the current study presents the consequences of having low early academic performance, there is also a silver lining to be found in the current study because policy makers and educators alike can now have a better understanding of the extent early success impacts future student performances. This research shows that there is a need to focus on the foundational academic skills of students in order to overcome the inequities inherent in HIDOE public schools. Policy and pedagogical interventions at the early grades can also be formulated for students with Native Hawaiian ancestry so as to mitigate the disadvantages of culture and poverty on their future academic achievement in Hawaii's public schools.

The current research can also inform on the guidelines for future investigation on the impact of early academic achievement on future academic success. One future research could be using proficiency levels instead of standard scores for reading or math for the 2002 cohort. Future research on other cohort of students can similarly use predictive models that are based upon the underlying multilevel models developed in the current study. With technology and software improvements in the future, it would be possible to model a data structure with more than one level of cross-classification in order to study the cumulative impact of early success on future academic achievement.

Furthermore, with assessments now being developed that are vertically linked by the HIDOE, it would be possible to develop growth models for future cohort of students to examine the impact of early success on their future academic achievement.

References

Abbot, M.L., & Joireman, J. (2001, July). *The relationship among achievement. Low income and ethnicity across six groups of Washington State students* (Technical Report #1). Seattle, WA: The Washington School Research Center.

Aikens, N.L., & Barbarin, O.A. (2008). Socio-economic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*(2), 235-251.

Bodley, J. (1994). An anthropological perspective. *In cultural anthropology: Tribes, states, and the global system*. New York: McGraw-Hill.

Burkam, D., Ready, D., Lee, V., & LoGerfo, L. (2004). Social-class differences in summer learning between kindergarten and first grade: Model specification and estimation. *Sociology of Education, 77*(1), 1-31.

Chard, D., Stoolmiller, M., Harn, B., Wanzek, J., Vaughn, S., & Linan-Thompson, S. (2008). Predicting reading success in a multilevel school-wide reading model: A retrospective analysis. *Journal of Learning Disabilities, 41*(2), 174-188.

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the early childhood longitudinal study (ECLS) Kindergarten to first grade sample. *Journal of Educational Psychology, 98*(3), 489-507.

Chatterji, M. (2005). Achievement gaps and correlates of early mathematics achievement: Evidence from the ECLS K–first grade sample. *Education Policy Analysis Archives, 13*(46).

Department of Education (2008). *Foundations for Success: The Final Report of the*

National Mathematics Advisory Panel. US Department of Education, Washington D.C. Retrieved from ERIC database. (ED500486).

Driessen, G. (2002). School composition and achievement in primary education: A large-scale multilevel approach. *Studies in Educational Evaluation, 28*(4), 347-368.

Garcia, D. (2007). Mixed messages: American Indian achievement before and since the implementation of No Child Left Behind. *Journal of American Indian Education, 46*(3), 136-154.

Gersten, R., Jordan, N.C., & Flojo, J.R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disability, 38,* 293-304.

Hanushek, E. A., & Rivkin, S. G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of Policy Analysis and Management, 28*(3), 366-393.

Harkreader, S., & Weathersby, J. (1998). *Staff development and student achievement: Making the connection in Georgia schools*. Atlanta, GA: The Council for School Performance.

Hawaii Department of Education. (2008). *Consolidated state application accountability workbook for state grants under Title IX, Part C, Section 9302 of the Elementary and Secondary Education Act (Public Law 107-110).* Unpublished manuscript.

Hawaii Department of Education. (2003b). *Teacher's guide for interpreting the Hawaii Content and Performance Standards* (2nd ed.) Honolulu: Hawaii Department of Education.

Hecht, S., Torgesen, J.K., Wagner R.K., & Rashotte, C.A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology, 79*(2), 192-227.

Hernandez, D.J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation.* Baltimore, MD. Annie E. Casey Foundation.

Hungi, N. (2008). Examining differences in mathematics and reading achievement among Grade 5 pupils in Vietnam. *Studies in Educational Evaluation, 34*(3), 155-164.

Kana'iaupuni, S. M., & Ishibashi, K. (2003). *Left behind: The status of Hawaiian students in Hawai'i public schools.* http://www.ksbe.edu/pase/pdf/reports/k-12/02_03_13.pdf. Retrieved on May 22, 2009.

Kieffer, M. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology, 100*(4), 851-868.

Kilian, L., & Kagen, E. (1981, April). *The long term effects of the ESEA Title I reading program on reading achievement.* Paper presented at the Annual Meeting of the American Educational Research Association. Los Angeles, CA. Retrieved from ERIC database. (ED198949).

Liu, X., & O'Connell, A. (2005, October). *Using multilevel modeling for change to assess early children's reading growth over time.* Paper presented at the Annual Meeting of the Northeastern Educational Research Association. Kerhonkson, NY. Retrieved from ERIC database. (ED501279).

Luyten, H., Peschar, J., & Coe, R. (2008). Effects of schooling on reading performance, reading engagement, and reading activities of 15-year-olds in England. *American Educational Research Journal, 45*(2), 319-342.

McNiece, R., Bidgood, P., & Soan, P. (2004). An investigation into using national longitudinal studies to examine trends in educational achievement and development. *Educational Research, 46*(2), 119-136.

Milazzo, P. (1981). *Method for analyzing district level data bases to identify learning opportunity risks.* Retrieved from ERIC database. (ED251418).

Nochi, G. (2008). *Differential effects of poverty on Hawaii students' achievement in reading and mathematics*. Unpublished doctoral dissertation, University of Hawaii.

Patterson, C.J., Kupersmidt, J.B., & Valden, N.A. (1990). Income level, gender, ethnicity, and household composition as predictors of children's school-based competence. *Child Development, 61,* 485-494.

Peng, S.S., & Wright, D. (1994). Explanation of academic achievement of Asian-American students. *Journal of Educational Research, 87*(6), 346-352.

Pratt, A. C., & Brady, S. (1988). Relation of phonological awareness to reading disability in children and adults. *Journal of Educational Psychology, 80*(3), 319-323.

Princiotta, D., & Hausken, E. G. (2006). *Fifth grade: Findings from the fifth grade follow-up of the early childhood longitudinal study, Kindergarten class of 1998 to 1999 (ECLS-K).* National Center for Education Statistics. Retrieved from ERIC database. (ED491189).

Pustjens, H., Van de gaer, E., Van Damme, J., Onghena, P., & Van Landeghem, G. (2007). The short-term and the long-term effect of primary schools and classes on mathematics and language achievement scores. *British Educational Research Journal, 33*(3), 419-440.

Rathbun, A., West, J. & Hausken, E.G. (2004). *From kindergarten through third grade children's beginning school experiences.* NCES 2004 to 2007. US Department of Education. Retrieved from ERIC database. (ED483078).

Rugutt, J., & Ellett, C. (2001). *A study of students' academic change in mathematics and language achievement: A multilevel structural equation model (MSEM) approach*. Retrieved from ERIC database. (ED467835).

Rugutt, J., Ellett, C., & Kennedy, E. (2002). *A study of students' academic change in mathematics achievement: A case for African American students.* Retrieved from ERIC database. (ED468687).

Rugutt, J. (2001). *A study of individual patterns of longitudinal academic change: Exploring the structural equation modeling (SEM) and hierarchical linear modeling (HLM).* Retrieved from ERIC database. (ED467836).

Saturnelli, A.M., & Repa, J.T. (1995, April). *Alternative forms of assessment in elementary science: The interactive effects of reading, race, economic level and the elementary science specialist on hands-on and multiple-choice assessment of science process skills*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA. Retrieved from ERIC database. (ED387519).

Takanishi, S. (2005). *A multilevel analysis estimating schools' likelihood of meeting*

NCLB academic targets: A comparison of two models of effectiveness. Unpublished doctoral dissertation, University of Hawaii.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second-to fifth-grade children. *Scientific Studies of Reading, 1*(2), 161-185.

Uyeno, R., & Zhang, S. (2007, February). *What the Hawaii State Assessment at Grade 3 tells us about academic achievement at Grade 7*. Paper presented at the 29th Annual Meeting of the Hawaii Educational Research Association, Honolulu, HI.

Van der Berg, S. (2008). How effective are poor schools? Poverty and educational outcomes in South Africa. *Studies in Educational Evaluation, 34*(3), 145-154.

Williams, M.M. (1972, September). *Race, poverty and educational achievement in an urban environment*. Paper presented at the American Psychological Association Annual Convention, Honolulu, HI. Retrieved from ERIC database. (ED070797).

Wong, K.K., & Alkins, K.F. (1999, April). *Toward systemic reform in high poverty schools: A comparative analysis of two large school districts*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada. Retrieved from ERIC database. (ED433398).

Zhang, S. (2009). The impact of economic disadvantage on academic achievement in learning in Hawaii: A multilevel analysis. *International Journal of Learning*, *16*(7), 367-382.