Asian Americans in Educational Research:

The Use of Disaggregated Racial and Ethnic Subgroup Data

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Mao Thao Jacobson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Dr. Frances Lawrenz, Advisor

January 2019

Acknowledgements

I have come this far only because of those who walked beside me when I needed motivation, backed me up when I needed strength, and took the lead when I felt lost. It is important to acknowledge those who have been part of this journey with me.

My husband has made sure I never felt alone in this endeavor. Lee, your fierce love gives me strength and pride to be my authentic self. Thank you for believing in me, for taking on more than your share at home, and for taking care of our family.

My advisor and committee members have supported my professional and academic growth. Dr. Frances Lawrenz, there were many times you would say aloud, "What would I tell my daughter?," before responding to my questions. Thank you for advising from a place of love, yet being firm and reminding me I had work to finish. Dr. Jean King, thank you for inspiring me to be a more mindful evaluator. Dr. Mark Davison, your standards for academic and research ethics has made me a stronger evaluator and researcher, thank you. Dr. Andrew Zieffler, I am grateful for your support in finding solutions to my questions and problems, thank you.

I am forever thankful to the faculty and the administrative staff of the Quantitative Methods in Education program. To the faculty, thank you for pushing me to go above and beyond what I thought I was capable of. To the administrative team, especially Sharon and Lori, thank you for keeping me on track.

I am so appreciative to have a strong network of family and friends. Rau kuv niam thiab txiv, kuv yawg thiab pog, thiab kuv niam tais, ua tsaug rau nej txoj kev hlub thiab txoj kev txhawb nqa. Yog tsis muaj nej, tse kuv yuav tsis kawm kom tiav. Rodney, Alice, and Grandma Alice, thank you for your loving support over the years. To my siblings and extended family, thank you for your words of encouragement. Thanks to my friends who kept me sane. Sue, Johnny, and Youa, thank for keeping my son company so I could have time to write. Soua Yee, thank you for being there to listen and pick me up when I needed it most. To my QME peers—Amy Grack Nelson, Lija Greenseid, Alison Phillips, Katherine Edwards, Mario Moreno, Elizabeth Fry, Anica Bowe, Christopher Moore, Ann Betzner, and Laura Le—thank you for being part of this roller coaster journey and for the pep talks along the way! Amy, it means the world to me to have gone through this with you. Thank you for always reminding me of the light at the end of the (long) tunnel.

I could not have completed this work without the support of my colleagues, who were generous with their time to listen and provide their perspectives. Thanks to Stacey Gray Akyea, John Lindner, Pang Nhia Yang, Holly Miller, Joe Munnich, and Marian Heinrichs. Your guidance and support through the very end of this process means a lot to me. I must also acknowledge Ann Hironaka, Cheryl Holm-Hansen, Laura Martel-Kelly for writing my recommendation letters years ago. I would not be here without your support, thank you. Additionally, I would like to thank the National Center of Education Statistics and the Minnesota Department of Education for providing access to the data sets used in this research.

Above all, I owe thanks to my son. Mi tub, I will cherish forever the way your face lit up after I told you I got a ‘thumbs up.” Thank you for being patient with me and for inspiring me to be a better human being in every way possible. Kuv txoj kev hlub rau koj tsis muaj hnub kawg.

Dedication

For my parents and grandparents, your unwavering love and support has made all the difference in accomplishing this doctoral degree. I can only imagine what it is like to give up everything you have, live in fear, and risk your lives for a chance at freedom. Your journey of survival gives me strength and inspiration to persevere no matter what obstacles I face.

For the Hmong who perished during the Secret War in Laos, your fateful sacrifice in search of a safe and free home were not made in vain.

Abstract

The Model Minority stereotype of Asian American students is widespread. It suggests Asian Americans (AA) are a homogeneous group and achieve high levels of academic, social, and economic success. Aggregating data on AA students in educational research have perpetuated this stereotype by showing that AA students outperform their racial counterparts. There is growing advocacy to disaggregate data on AA students by racial and ethnic subgroup because achievement gaps are concealed when AA are grouped together. In particular, Southeast Asian subgroups typically have lower levels of academic achievement compared to other AA subgroups and are at risk of being overlooked for supports and opportunities. This study reviews current research on the academic achievement of AA students, historical changes in how AA have been classified, the impact and validity of using aggregated data on AA students, and the ways disaggregated data on AA students have been used. With data from the High School Longitudinal Study of 2009 and Minnesota student assessment data, this study explores: 1) the bimodality/multimodality of math test scores within racial group for Asian, Black, and White students; 2) mean differences in math achievement across Asian subgroups, and 3) the use of an aggregated Asian group compared to disaggregated Asian subgroups to examine math achievement growth during high school across race. Findings are used to provide practical implications for educational researchers, evaluators, policy makers, and decision makers to consider when using disaggregated data.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Asian Americans (AA) have been plagued with the "model minority" stereotype for over 50 years. The stereotype was conceived in 1966 when two articles highlighted Japanese and Chinese Americans as examples of how a marginalized minority group has risen above social prejudice and historical injustice to be successful in the United States (U.S.). In the first article, published in the *New York Times*, sociologist William Petersen (1966) praised Japanese Americans in persevering against racial discrimination, the denial of rights, and confinement in concentration camps during World War II. Minority groups enduring such discrimination would have likely produced what he called "problem minorities" (p. 20). Petersen attributed the success to strong cultural beliefs, family values, and work ethics. In the second article, published in *U.S. News and World Report*, authors described Chinese Americans as a "model of self- respect and achievement" ("Success Story," 1966, p. 73). The authors highlighted low crime rates in Chinatown enclaves and the group's discipline and hard work in education and the workforce.

Critics of the model minority stereotype argue it was conceived to counter the Civil Rights Movement during the 1960s (Lee, 1996; Museus, 2014; Suzuki, 1977; Wang, 2008; Yu, 2006). The two articles were not written coincidentally, but were a reflection of the political strife around racial equality at the time. The premise of setting up AA as the model minority asserts that it is possible for marginalized racial minorities to be successful independent of federal assistance and intervention. Sentiments of that argument can be inferred from the second article, as the headline reads:

> At a time when Americans are awash in worry over the plight of racial minorities – one such minority, the nations' 300,000 Chinese-Americans, is winning wealth and respect by dint of its own hard work. … Still being taught in Chinatown is the old idea that people should depend on their own efforts – not a welfare check – in order to reach America's "promised land." ("Success Story," 1966, p. 73)

The stereotype persists regardless of the political propaganda behind the image of AA as the model minority. Mainstream media has continued to tout the success of AA, describing them as "whiz kids" and "super minorities" (Hartlep, 2014). A Pew Research Center (2013) report entitled, "*The Rise of Asian Americans*," highlighted the economic and educational success of AA. The report overview opened with:

> Asian Americans are the highest-income, best-educated and fastest-growing racial group in the United States. They are more satisfied than the general public with their lives, finances and the direction of the country, and they place more value than other Americans do on marriage, parenthood, hard work, and career success. (Pew Research Center, 2013, p. 1)

The report received criticism from the AA community for continuing to incite the rhetoric of the model minority stereotype (Hing, 2012; Lee & Zhou, 2015; Museus, 2014; Watanabe, 2015). Many AA community organizations commended the Pew Research Center for its efforts to further current research about AA, but expressed concerns about the overgeneralization of AA and use of aggregated data that conceal disparities across AA subgroups (see https://foundasian.org/2012/06/). While the report recognized the diversity within the AA community and identified economic disparities among AA subgroups, the headlines and framing of the findings were reminiscent of prior publications and media articles spreading the perception of AA as the model minority. The news media picked up two primary messages from the report: (1) the success of AA above and beyond other racial and ethnic groups, and (2) the increase in Asian immigrants surpassing that of Hispanic immigrants (Guey & Lubin, 2013; Martin, 2012; McCarthy, 2012; Munroe, n.d.; Siegel, 2012). Whether or not it was a response to the criticism, the report was revised nearly a year after its initial release to include additional U.S. Census data for smaller AA subgroups (e.g., Bangladeshi, Cambodian, Hmong, Malaysian, Lao, Pakistani, Thai, Indonesian, Sri Lankan) (Pew Research Center, 2013).

There is growing concern within the AA community around research and data on AA and how they perpetuate the model minority stereotype, often referred to as the "Model Minority Myth" (MMM). Museus and Kiang (2009) define the MMM as "the notion that Asian Americans achieve universal and unparalleled academic and occupational success" (p. 6). Many AA are successful and their accomplishments should not be discounted, however, that is not the reality for all AA subgroups. The MMM implies that AA are a homogeneous group who are well off and not in need of support (Lee, 1996; Museus & Kiang, 2009). Those assumptions are made from research using aggregate data on AA that disregard differences in culture, language, and historical background across subgroups that impact the social and economic circumstances of families, as well as academic outcomes for children. Moreover, researchers infrequently identify key characteristics of AA study participants that have been shown to be important predictors of educational and economic achievement, such as subgroup, immigrant generational status, and English language proficiency (Barrett, Barile, Malm, & Weaver, 2012; Harris, Jamison, & Trujillo, 2008; Lee & Zhou, 2015; Portes &

Rumbaut, 2001; Tamura, 2003; Tran & Birman, 2010; Zhang, 2001). Such details would give consumers of the research, like policy makers and program providers, a better contextual understanding of the findings. Yet data users are left to draw conclusions about AA overlooking existing disparities across subgroups. As a result, some AA subgroups are left invisible and struggle to achieve the high level of success shown in research with aggregated data and the MMM.

The widespread stereotype of AA as model minorities achieving high academic and financial success is a false representation of the realities of being AA in the U.S. While the stereotype was initiated in news media, it has continued to persist in educational research. Racial disparities research using aggregated data continue to show that AA students outperform students of all other racial groups. The research is flawed as aggregated data conceal significant disparities across AA subgroups. Southeast Asian students, in particular those with families who resettled in the U.S. as refugees, are facing inequitable access to needed educational opportunities and supports. There is growing advocacy and demand for research using disaggregated data on AA (National Commission on Asian American and Pacific Islander Research in Education (CARE), 2013). Along with it, there is a need for more consistency and guidance to using disaggregated data. This study focuses on the use and impact of aggregated and disaggregated data on Asian students in educational research and evaluation. It will examine math disparities within and across race groups for Asian students, White students, and Black students. Study findings will help provide practical implications for educational researchers and evaluators, as well as policy makers and decision makers.

Chapter 2: Review of Literature

Asian American Students and Academic Achievement

The educational research field is a prime example of how aggregated data have played into MMM. It has been well documented that AA students academically outperform other racial and ethnic groups. The most recent national report on high school transcripts indicate that AA high school graduates in 2009 had a higher grade point average than other groups, 3.26 compared to 3.09 for White, 2.84 for Hispanic, and 2.69 for Black students (Nord et al., 2011). Among a nationally representative sample of 4th, 8th, and 12th graders, the 2017 National Assessment of Educational Progress (NAEP) data show higher proportions of AA students are at or above proficiency in math and reading compared to other racial and ethnic groups (Snyder, 2018). Among the graduating class of 2017 who took the ACT, AA had the highest mean subject and composite scores in comparison to their other racial peers (ACT, 2017). The average composite score for Asian students was 24 compared to 22 for White, 19 for Hispanic, 18 for both Native Hawaiian and Pacific Islander (NHPI) and American Indian/Alaskan Native, and 17 for Black students. AA also have higher college enrollment and degree attainment rates compared to other racial groups (Snyder, 2018). In 2016, 67% of Asian young adults (ages 18 to 24 years old) were enrolled in a degree-granting post-secondary institution compared to 45% of Whites, 36% of Blacks, 36% of Hispanics, 36% of Pacific Islanders, and 22% of American Indians/Alaskan Natives. Furthermore, AA (ages 25 and older) have had higher rates of attaining at least a Bachelor’s degree than their racial counterparts for more than 20 years. In 2017, the attainment rate for a bachelor’s degree or higher was 55% for Asian adults 25 years and older, while the attainment rate was 38%

for Whites, 25% for Pacific Islanders, 24% for Blacks, 17% for Hispanics, and 21% of American Indian/Alaskan Natives.

Aggregated and nationally representative data on AA support the MMM, however, a different story unfolds when data are disaggregated by subgroup. Many researchers and advocacy groups have used the U.S. Census data to show disparities across AA subgroups and dispel the MMM. Southeast Asian subgroups whose families immigrated as war refugees (e.g., Burmese, Cambodian, Hmong, Lao, Khmer, Vietnamese) typically have a large youth population, higher poverty rates, and lower educational attainment levels than other Asian subgroups (CARE, 2010; CARE, 2013; Ngo & Lee, 2007; Pak, Maramba, Hernandez, 2014; Yang, 2004). According to the 2017 American Community Survey 1-year estimates (U.S. Census Bureau, 2017, Table S0201):

- Children under 18 years old made up 20% of the overall Asian population, but ranged between 10% to 35% across Asian subgroups. Hmong and Burmese subgroups had the largest population of youth under 18 years old (35% for both groups).
- The poverty rate for Asian families with minor children was 9%, but varied widely across Asian subgroups where Burmese families have the highest poverty rate at 33% and Asian Indians have the lowest poverty rate at 4%.
- Among the Asian adult population (ages 25 years old or older), 54% had an educational attainment level of at least some college or an associate's degree. The proportion of adults with this level of educational attainment ranged between 32% and 89% across Asian subgroups, where Burmese (32%), Lao

> (43%), Cambodian (43%), Hmong (53%), and Vietnamese (53%) subgroups had lower proportions than other Asian subgroups.

Furthermore, the Pew Research Center (2018) reports that Asians are the most economically divided group in the U.S. Using an income measure that adjusts for household size, the income inequity ratio between the top and bottom 10% of the Asian population increased from 6.1 in 1970 to 10.7 in 2016. This was the highest percent increase in ratios across race and ethnic groups. A ratio of 10.7 means that those in the top 10% earned 10.7 times more than those in the bottom 10% among Asians in 2016. In the same year, the income inequity ratio was 9.8 for Blacks and 7.8 for both Whites and Hispanics. An explanation provided for this inequity was immigration and workforce patterns. Between 1970 and 2016, immigration policies resulted in an influx of highly educated Asian immigrants to fill high-skill job positions but also the arrival of Asian war refugees with lower education levels.

Researchers have further examined the heterogeneity within the AA community and its impact on educational research. Teranishi (2010) argues there is a great deal of variation in standardized test scores among AA and Pacific Islander students compared to other race groups that go unnoticed, where the spread of test scores for AA and Pacific Islander students vary from the mean score. In examining the Scholastic Aptitude Test (SAT) scores in 2004, he highlights that the distribution of math scores for AA and Pacific Islander students showing evidence of bimodality. Scores for White students were normally distributed and centered at the mid-point of scores between 500 and 549, where nearly 19% of students scored in this range. Scores for Asian students were more uniform across mid-to-high score ranges, however, the percentage of students scoring at the mid-point of 500 and 549 decreased slightly from 14% to 12%, then increasing back to 14%. Teranishi (2010) argues that the performance of AA and Pacific Islander students on the SAT is confounded with that of Asian international students who apply to selective higher educational institutions in the U.S., noting that the College Board groups together "Asian, Asian Americans, or Pacific Islander" (p. 112). In 2006, almost half of all international students were from five Asian countries, including India, China, Korea, Japan, and Taiwan (Open Doors, 2007 as cited in Teranishi, 2010). Further, he posits that parental education, parental income, and English proficiency are correlated with Asian students' verbal scores and are likely contributing to the bimodality of math scores, where students from families with low parental educational attainment, low parental income, and limited English proficiency were more likely to have lower math scores than their counterparts.

Reeves and Halikias (2017) report on more recent 2015 SAT scores and include a similar analysis as Teranishi (2010) examining the percentage of students across scores ranges of 50 points by race. Reeves and Halikias (2017) comment on the relatively normal distribution of scores for White students, where 19% of students had scores between 500-550. Scores for Asian students were negatively distributed peaking at scores between 650 and 700 with 15% of the students; thus, not showing evidence for bimodality. What is striking is the percentage of Asian students in the highest score range increased considerably between data presented in these two studies. In Teranishi's (2010) study, about 5% of Asian students scored between 750 and 800 in 2004, while Reeves and Halikias (2017) report about 14% of Asian students did so in 2015. As a limitation, Reeves and Halikias (2017) state race gaps in SAT scores could be even wider if there was not a ceiling on SAT scores, meaning Asian students could very well achieve high scores beyond what is allowed on the SAT. But, they did not address the concerns highlighted by Teranishi (2010) regarding Asian international students. In 2016-17, the first year 2015 SAT test-takers would be enrolled in college, the majority of international students (about 68%) were from Asian countries (Institute of International Education, 2018a). China (33%) and India (18%) continue to be the top countries where international students are from. Further, the top three fields of study among international students in 2016-17 were engineering, business and management, and math and computer science (Institute of International Education, 2018b). These fields require high levels of math skills and achievement and thus, could very well contribute to explaining the high SAT scores among Asian students.

In an effort to detail the heterogeneity that exists with the Asian population, Hartlep, Morgan, and Hodge (2015) examined the extent to which there were underlying subpopulations among Asian students in 10th grade from the Educational Longitudinal Study (ELS) of 2002. They found three of subgroups within a sample of 1,070 Asian students using finite mixture modeling, a method that classifies subgroups based on similarities in demographics. The three subgroups were similar in size, with subgroup 1 consisting of 29% of the sample, subgroup 2 with 33% of the sample, and subgroup 3 with 38% of the sample. Subgroups differed primarily by home language, parental educational attainment, parental educational aspirations, urbanicity, U.S. region, socioeconomic background, and average composite test score assessing math and reading. The first subgroup consisted largely of students with a West or South Asian home language (68%), with parents with low levels of educational attainment of a high school diploma/GED or less (94% of mothers and 82% of fathers), with high parental educational aspirations for them to earn a 4-year college degree (47%), and lived in urban areas (57%) in Western U.S. (48%) or the Midwest (23%). Compared to the other two subgroups, this group had the lowest family socioeconomic index ($\bar{x}$ = -1.0, *SD* = 0.4) and the lowest average standardized composite test score ($\bar{x}$ = 48.1, *SD* = 9.0).

Subgroup 2 had high proportions of students with an English (45%) or West or South Asian home language (37%), high levels of parental educational attainment of at least a 4-year college degree (88% of mothers and 94% of fathers), with very high parental educational aspirations for them to earn a doctoral degree (43%), and lived in suburban areas (62%) in Western (46%) or Southern (22%) U.S. This subgroup had the highest family socioeconomic index ($\bar{x}$ = 1.0, *SD* = 0.3) and average test score composite ($\bar{x}$ = 59.5, *SD* = 8.9) compared to the other two subgroups. The third and last subgroup included high proportions of students with a West or South Asian home language (52%), with varied parental educational attainment from a high school diploma/GED to a 4-year college degree (34% of mothers had a high school diploma/GED, while 55% of mothers had up to a 4-year degree; 22% of fathers had a high school diploma/GED and 30% had a 4-year college degree), with high parental expectations to achieve a 4-year college degree (40%), who lived in suburban areas (54%) in Western U.S. (44%). Subgroup 3 had average family socioeconomic index ($\bar{x}$ = 0.0, *SD* = 0.3) and an average standardized composite test score ($\bar{x}$ = 52.7, *SD* = 8.8) between that of the other two groups. Hartlep, Morgan, and Hodge (2015) conclude that their findings disproves the MMM by showing that the Asian population is indeed quite heterogeneous.

Research studies on AA have also demonstrated how aggregated data can be misleading at all levels of education. Yang's (2013) study using data from the Early

Childhood Longitudinal Study—Kindergarten Class 1998-1999 showed achievement gaps in reading and math across AA subgroups in the first grade. Southeast Asian and Filipino students had significantly lower mean reading and math assessment scores than other AA subgroups. In math, Asian Indian students had significantly higher mean scores than Filipino students by 7.2 points, other Asian students also by 7.2 points, and Southeast Asian students by 6.5 points. Greater disparities existed in mean reading scores, where Southeast Asian students scored 18.3 points lower than Asian Indian students on average. When Yang (2013) controlled for school (urbanicity; funding type, and size) and family (parental marital status, family poverty level, and primary home language) factors, disparities across Asian subgroups continued to exist, but mean score differences were not as drastic. In math, only Filipino students had a significantly lower mean math score than Asian Indian students by 5.2 points after controlling for school and family demographics. In reading, however, four subgroups performed lower than Asian Indians after accounting for school and family factors: Filipino students by 5.9 points, Japanese students by 6.0 points, Southeast Asian students by 10.4 points, and other Asian students by 10.7 points.

In an educational policy report, the Council on Asian Pacific Minnesotans (2012) showed lower percentages of Burmese, Cambodian, Hmong, and Lao students meeting the Minnesota student assessment reading and math standards compared to other AA subgroups, as well as the aggregated Asian group. Aggregated data show 55% of Asian students met proficiency standards in math, but only 47% of Hmong students and 17% of Burmese students did so. Additionally, 55% of Asian students as an aggregate group were proficient in reading, but only 37% of Hmong students and 12% of Burmese students were proficient—much lower percentages than White (81%), American Indian (54%), Hispanic or Latino (53%) and Black or African American (52%) students.

Pang, Han, and Pang (2011) examined California standardized assessment data for 7th graders between 2003 and 2008. They found AA and Pacific Islanders, as an aggregate group, had a significantly higher mean math achievement score than their White peers with a difference of 4.5 points. When the data were disaggregated for 13 Asian subgroups, seven of those groups had a significantly lower mean math score than White students, including Filipino, Guamanian, Native Hawaiians, Cambodian, other Pacific Islanders, and Lao students. Mean differences ranged from 0.6 to 13.8 points across these Asian subgroups compared to White students. Asian students, as an aggregate group, had a significantly lower mean reading score than White students by 1.4 points. However, when using disaggregated Asian subgroups, three groups had a higher mean reading score: Japanese students by 4.8 points, Chinese students by 4.7 points, and Korean students by 3.1 points. The remaining Asian subgroups had lower mean reading scores than White students ranging from 0.5 to 16.7 points. Across all groups, Cambodian, Lao, other Pacific Islander, and Samoan students had the lowest math and reading scores. Pang et al. (2011) further chose to compare math and reading scores across Chinese students, Samoan students, and White students while accounting for sex, free or reduced lunch status, and parents' education level. Chinese students had significantly higher mean scores in both math and reading than White students, while Samoan students had significantly lower mean scores than White students.

Extending on the research completed by Pang et al. (2011), Lee et al. (2017) used the same California assessment data from 2003 to 2008 to assess reading and math disparities across four student groups: 1) Vietnamese students, 2) Lao or Cambodian students, 3) White students, and 4) Black students. Lao and Cambodian students were combined as one group because of their small sample sizes. Disparities were examined while accounting for free or reduced-price lunch status and parents' educational level separately. Among students who were eligible for free or reduced-price lunch, Lao/Cambodian students and Black students had lower reading and math mean achievement scores compared to White students, 5.1 points and 9.8 points respectively in reading and 1.2 points and 10.9 points respectively in math. Vietnamese students had higher scores than White students on average by 4.9 points in reading and 12.8 points in math. Across four categories of parental educational level (less than high school graduate, high school graduate, some college, and college graduate or graduate school), students whose parents had low levels of education typically had lower mean reading and math scores than those with highly educated parents. In both reading and math, Lao/Cambodian students and Black students typically scored lower than White students, whereas Vietnamese students scored higher than White students on average.

Her (2014) examined results from the Early Assessment Program, where all 11th graders in California take assessments to measure readiness for college-level math and English courses at California State Universities (CSU). Students who are deemed not ready for college are required to take placement exams, which determine whether remedial courses are necessary during their freshman year at CSU. Aggregated data showed that fewer Asian students were "not ready" for college-level courses compared to White students across three years of data (2011-2013). In 2013, 40% of Asian students and 49% of White students were "not ready" for college-level English classes, while 16%

of Asian and 32% of White students were “not ready” for college-level math courses. However, disaggregated data revealed that much higher percentages of some Southeast Asian groups (Cambodian, Hmong, and Lao) were “not ready” for college-level courses compared to other Asian subgroups, White students, and the aggregated Asian group. Hmong students, in particular, had the highest percentage of being “not ready” for college coursework—81% were “not ready” for college-level English courses, while 40% were “not ready” for college-level math courses in 2013.

Teranishi (2010) finds subgroup differences in college enrollment patterns among a sample of Asian students across the U.S. in 1997. Across five Asian subgroups, Filipino students and Southeast Asian students were more likely to enroll in a two-year college or a public 4-year college, while Chinese students, Korean students, and Japanese students were more likely to attend private colleges. After accounting for high school experiences and college choice factors, Chinese students and Korean students were significantly more likely to attend a selective college, while Filipino students and Japanese students were significantly less likely to attend a selective college. Results did not indicate that Southeast Asian students were more or less likely to attend a selective college, meaning enrollment in a selective college varied too much within this subgroup. Parental income and educational attainment have a positive association with students’ enrollment in a selective college, where students whose parents had higher levels of income and education were more likely to attend a selective college.

Immigration patterns and histories provide context for the wide range of educational and socioeconomic backgrounds across AA subgroups. Immigration of individuals from East and South Asia go back much earlier in U.S. history than immigration of those from Southeast Asian countries. Immigrants from China, Japan, the Philippines, Korea, and India have commonly come to the U.S. in search of work and education. The Immigration Act of 1965 was pivotal as it lifted the bans on immigration from Asia (Tamura, 2003; Lee & Zhou, 2015). Early Asian immigrants before this time typically came to fill manual laborer positions and were faced with high poverty and racial discrimination, such as the denial of naturalized citizenship, segregated schools, the Chinese Exclusion Act of 1882 barring immigrants from China, the Immigration Act of 1924 barring immigrants from Asia with the exception of the Philippines[1], and the internment of Japanese Americans during World War II.

The intent of the Immigration Act of 1965 was an effort to reunify families for immigrants from Europe and build the U.S. skilled workforce (Lee & Zhou, 2015; Tamura, 2003; Zong & Batalova, 2017b). Yet there was an influx of immigrants from Asian countries. Immigration patterns from Asia post-1965 are characterized by an influx of highly educated and skilled Asian immigrants, as well as the resettlement of war refugees with mixed levels of educational attainment and professional work skills (Lee & Zhou, 2015; Tamura, 2003). Preference was given to family members of U.S. citizens and skilled workers with advanced degrees (Center for Immigration Studies, 1995; U.S. Citizenship and Immigration Services, n.d.). When the list of preferences was revised in 1990, it specified priority on those with "extraordinary" or "exceptional" abilities (U.S. Citizenship and Immigration Services, n.d.). Low skilled workers were also targeted based on the demand, but they were of lower priority. Those filling the demand for highly

[1] Filipinos were considered U.S. nationals as the Philippines was a U.S. colony until 1946 (see https://en.wikipedia.org/wiki/United_States_territorial_acquisitions#Philippines.

skilled workers were primarily from mainland China, but also from Korea, India, and the Philippines (Tamura, 2003; Zong & Boatalova, 2017b).

There are three primary waves of war refugees during this time as a result of the Vietnam War, but also civil wars in Laos, Cambodia, and Myanmar (formally known as Burma) (Kula & Paik, 2017; Lee, 2015; Lee & Zhou, 2015; Ngo & Lee, 2007; Portes & Rumbaut, 2001). Refugees from those countries included Burmese, Cambodians, ethnic Chinese, Hmong, Karen, Khmer, Lao, and Vietnamese subgroups. The first wave between 1975 and 1978 consisted of educated professionals and those who worked with the U.S. Central Intelligence Agency (CIA). The vast majority of refugees during this wave were from Vietnam, who typically lived in urban areas, had some English language skills, and had professional and business work experience (Kula & Paik, 2017). The second wave of refugees came between 1979 and 1982 and were primarily admitted under family reunification policies, thus many were family members of the first wave of refugees. Most refugees continued to be from Vietnam, but there were more refugees from Laos and Cambodia in this wave than the first (Kula & Paik, 2017). These refugees typically had limited or no English proficiency and were farmers or agrarians from rural areas with limited access to education. The third wave of Southeast Asian refugees, 1982 to the present, include those who had been living in refugee camps in the aftermath of the Vietnam War. These refugees also had limited English proficiency and lacked formal education. In the mid-2000s, there was a spike in Hmong refugees, originally from Laos, who were resettled after the closing of the Wat Tham Krabok refugee camp in Thailand (Kula & Paik, 2017). Throughout the 2000s, there was also a rise in Burmese refugees as result of the civil war in Myanmar (Tandon, 2016). Between 2007 and 2017, Burmese refugees were the largest group admitted into the U.S., making up 23% of all 708,354 refugees (Zong & Batalova, 2017a).

Research using disaggregated data exemplifies the demand for a better understanding of the educational needs of AA students, especially Southeast Asian students. There has been slow progress to move towards collecting and analyzing disaggregated data in research. This review of literature will explore the methodological issues in conducting educational research on AA. The review begins with a discussion of how AA have been historically grouped and categorized. Next, an overview of the issues of aggregating data on AA and an examination of the validity of inferences are provided, followed by a summary of current efforts to disaggregate data on AA.

The Classification of Asian Americans as a Racial Group

It is important to consider the historical classification of AA as a racial group before delving into issues around aggregated data. The Asian race category has experienced changes impacting the consistency of data on AA over time.

The U.S. Census has an extensive history of collecting data on race and ethnicity as far back as 1790 (Humes & Hogan, 2009; Pratt, Hixson, & Jones, 2015). The U.S. Census first collected data on AA on a national level with a single "Chinese" group in 1870. Prior to that time, the "Chinese" race category had only been used in California where the majority of Chinese migrant workers resided. The most recent 2010 Census looks very different. It lists six Asian subgroups and three NHPI subgroups. There is also a write-in option to identify a subgroup not listed. The 2010 Census data are available for 25 detailed Asian subgroups and 20 detailed NHPI subgroups (Hixson, Hepler, & Kim, 2012; Hoeffel, Rastogi, Kim, & Shahid, 2012).

Changes in data collection for the census reflect immigration history. In the late 1800s, growth of Chinese and Japanese workers prompted additions to the race categories in the census (Humes & Hogan, 2009). In 1910, Koreans, Filipinos, and Asian Indians were grouped into an "other" category. Asian Indians are the only group to ever be categorized according to a religion in the U.S. Census. The term "Hindus" was used on the census between 1920 and 1940, while "Asian Indian" was first used in the 1980 Census. In addition, Asian Indians were not always categorized as Asian. Historically, those from the Indian sub-continent were initially identified as White. The following explanation was provided in the 1910 census:

> Pure-blood Hindus belong ethnically to the Caucasian or white race and in several instances have been officially declared to be white by the United States courts in naturalization proceedings. In the United States, however, the popular conception of the term white is doubtless largely determined by the fact that the whites in this country are almost exclusively Caucasians of European origin and in view of the fact that the Hindus, whether pure-blood or not, represent a civilization distinctly different from that of Europe, it was thought proper to classify them with non-white Asiatics (U.S. Census Bureau 1913, p. 126 as cited in Humes & Hogan, 2009, p. 115).

By the 1920 Census, Asian Indians (referred to as Hindus), Koreans, and Filipinos were finally listed as separate Asian subgroups. Vietnamese were added in the 1980 Census after the first wave of Southeast Asian refugees as a result of the Vietnam War. These six subgroups made up nearly 90% of the Asian population in 2010 (Pew Research Center, 2013).

The collection and reporting of federal data on race and ethnicity are guided by standards established by the Office of Management and Budget (OMB). In the mid-1970s, the OMB was charged with developing standard race and ethnicity categories for federal reporting after the U.S. Commission on Civil Rights identified inconsistencies in categories used across federal agencies (Humes & Hogan, 2009). As a result, the Statistical Policy Directive 15 was issued in 1997 by the OMB (OMB, 1997). It provided a set of standard definitions for four race categories (American Indian or Alaskan Native, Asian and Pacific Islander, Black, and White) and one ethnicity category (Hispanic). The following is the definition for the "Asian or Pacific Islander" category:

> A person having origins in any of the original peoples of Far East, Southeast Asian, the Indian subcontinent, or the Pacific Islands. This area includes, for example, China, India, Japan, Korea, the Philippines Islands, and Samoa (OMB 1978, p. 19269, as cited in Humes & Hogan 2009, p. 119).

In the early 1900s, the OMB began a review process to consider revisions to the Directive to better reflect the changing U.S. population (OMB, 1997). The OMB invited public comment and commissioned the Interagency Committee's Research Working Group on Racial and Ethnic Standards to conduct research and test several proposed changes regarding the categories and terminology used. The Directive was revised in 1997 based on the feedback received and research results.

The revised Directive resulted in several research implications regarding AA. The most publicized change of the revised Directive was the option to choose more than one race category. This change resulted in what looks like decreased population counts of some AA groups between the 1990 and 2000 census (Lai & Arguelles, 2003). For example, the Japanese population was estimated to be 847,562 in 1990, while the 2000 Census estimated a population of 796,700 persons identifying as Japanese only. When persons identifying as partially Japanese are included in the population estimate, it increases to 1,148,032. The change to recognize mixed race was a step towards greater accuracy, yet the comparability of data across years is unreliable, as it is not known what race category mixed AA respondents chose previously.

Another notable change in the revised Directive was the "Asian or Pacific Islander" category being split into two categories – "Asian" and "Native Hawaiian and Other Pacific Islander." The OMB implemented this change despite the Interagency Working Group recommendation against increasing the minimum number of race and ethnicity categories (OMB, 1997). The OMB specified the following definitions for the two categories:

> ***Asian***. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
>
> ***Native Hawaiian and Other Pacific Islander.*** A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands. (OMB 1978, p. 19,269, as cited in Humes & Hogan 2009, p. 119)

The split was motivated by Native Hawaiian advocates who successfully campaigned to be recognized as an indigenous group in the U.S. and argued that aggregated data using the "Asian and Pacific Islander" category did not accurately reflect the economic, health, and social conditions of the Native Hawaiian population (OMB, 1997). Initially, some advocates suggested being grouped in a “Native American” category along with American Indians and Alaskan Natives. Yet American Indian tribal groups opposed the idea as doing so would change the landscape of data on the American Indian population and complicate tribal legal status and self-government. These concerns weighed heavily on OMB’s decisions, especially as reporting of race and ethnicity data is used to enforce civil rights laws and allocate federal support and resources.

The split in categories was an effort for data to be reflective of the needs and conditions of the NHPI population, but the split has clouded educational research and data on AA. Asians and NHPI are considered separate races by the Directive, yet it is unknown the extent to which the two groups are consistently separated. Panapasa, Crabbe, and Kaholokula (2011) reviewed data collection and reporting procedures of 16 federal agencies, including administrative records and national surveys. All 16 agencies complied with OMB’s revised standard using separate race categories for Asians and NHPI, however, only one agency reported those data in detail – the U.S. Census Bureau. Two other agencies partially complied with the standards in data reporting, including the American Community survey and the Vital Fertility and Mortality Events. The remaining agencies either did not report the data at all due to small samples or reported the data using an aggregated category (e.g., AA and Pacific Islander, Other race).

The way AA have been classified has been inconsistent. Okazaki and Sue (1995) point out it is research subjects who decide which racial group they identify with, but the responsibility of choosing race and ethnicity categories for research lies in the hands of the researchers themselves. This dilemma also extends to the educational program evaluation community. Educational researchers and program evaluators face the challenge of deciding how AA and NHPI groups will be defined and included in their studies, if at all. Since their work is used to inform educational policy, it is important for them to be transparent in how they classify AA and justify their reasoning for grouping AA and NHPI. In addition, it is necessary for researchers and evaluators to provide details about the subgroup composition of AA samples. A lack of details about which AA subgroups are included could lead to misleading conclusions.

Asian Americans in the Aggregate: What's the Big Deal?

It is undeniable that AA, as an aggregated group, have higher levels of academic achievement than other racial groups. This finding has been consistent across many comparative studies and national statistics examining racial disparities in education over time. It is what the data show, and it has become common knowledge. On the surface, having greater academic achievement may appear to show an advantage for AA, however, there are underlying complexities that hinder understanding of the academic achievement among AA.

The adverse impacts of educational research on Asian Americans. Findings showing high levels of academic achievement among AA have had negative impacts on AA students. It has served as evidence for and perpetuated the MMM fostered in mainstream media. Some researchers have highlighted the adverse psychological effects of internalizing the model minority stereotype among AA students (Chu, 2002; Lee, 1994; Li, 2005; Tang, 2008; Thompson & Kiang, 2010; Wong & Halgin, 2006). Lee's (1994) ethnographic study found that both high- and low-achieving AA students tried to live up to the stereotype of Asians being high achievers. Students experienced anxiety, depression, and embarrassment from their attempt to meet the standards of the stereotype.

A student stated,

> They [Whites] will have stereotypes, like we're smart—They are so wrong; not everyone is smart. They expect you to be this and that, and when you're not—[*shakes her head*]. And sometimes you tend to be what they expect you to be, and you just lose your identity—just lose being yourself. Become part of what—what someone else want[s] you to be. And it's really awkward too! When you get bad grades, people look at your really strangely because you are sort of distorting the way they see an Asian. It makes you feel really awkward if you don't fit the stereotype. (Lee, 1994, p. 419)

Furthermore, there have been recent measurement instruments developed to assess the extent of internalizing the MMM and its relationship to students' psychological wellbeing (Shen, Wang, & Swanson, 2011; Yoo, Burrola, & Steger, 2010; Yoo, Miller, & Yip, 2015). Yoo, Miller, and Yip's (2015) measure, the Internalization of the Model Minority Myth Measure (IM-4), assesses perceptions of AA being more successful than other races in (1) achievement due to hard work ethics and (2) upward mobility due to lack of perceived racism. In their validation study, results showed the internalization of the MMM was related to stress from self-imposed academic expectations, as well as academic expectations from parents and teachers. When AA students had stronger beliefs of Asians being more academically successful than other racial groups, they experienced greater levels of stress from academic expectations.

The MMM also has adverse impacts within the classroom. When AA students fail to meet expectations of the MMM, they can be too embarrassed to seek academic help (Lee, 1994). AA students can also be very quiet in class (Chang & Au, 2014; Kim & Lee, 2014; Museus & Kiang, 2009; Park, 2000). Quietness is used as a coping strategy for distress and reflects Asian cultural values for respecting elders and authority figures. Teachers may mistake the quietness for shyness or good behavior—overlooking AA students who are struggling. In addition, teacher expectations can be shaped from research showing AA being more academically successful than other racial groups, as well as the MMM. When AA students' learning and behaviors deviate from what is expected, teachers may not know how to respond or provide support (Chang & Au, 2014; Li, 2005; Wright & Boun, 2011).

Another unfavorable impact of research on AA as high achievers is a lack of interest and investment in the education of AA students. The Asian population is largely ignored and invisible in education research and discourse (Museus, 2009; Museus, Maramba, & Teranishi, 2013). Museus (2009) found that only 1% of articles published in the top five peer-reviewed higher education journals between 1998 and 2008 focused on Asian American or Pacific Islander students. Because Asians typically do better academically than other racial groups, there is no disparity compared to White students. Many times, research studies and national reports on racial disparities do not even mention the Asian population at all. When AA students are of interest, the focus is more likely to be about what makes them so successful rather than underlying disparities. AA are not a group of high interest because it is assumed that AA students are doing well and are not in need of additional support and resources (Museus & Kiang, 2009). Museus and Chang (2009) maintain that a main challenge to educational research on AA is it can be exhausting to justify. They write:

> Researchers studying Asian Americans must not only directly debunk popular myths about universal success, but also justify their choice to include and study this population in the peer review process. We know scholars, for example, who in the past have been asked to explain why they included Asian American and Black participants (instead of Black and Latino students) in research on racial/ethnic minority undergraduates, whereas other researchers have been asked to justify their particular focus on particular Asian American ethnic subpopulations because reviewers perceived it as being too narrow. … Having to make such justifications can also become tiresome and dissuade scholars, especially those who are just starting their careers, from pursuing research on this population. (Museus & Chang, 2009, p. 97-98)

Those unaware of disparities between AA subgroups may not see the value in such research, however, it is important to continue to delve deeper into the accuracy behind findings on AA and academic success. Taking research findings in at face value and generalizing them to an entire population is a great disservice for AA subgroups that may need support.

Academic achievement findings favoring AA have also created an opportunity gap for AA students. Educational research guides policy and funding priorities at the federal, state, and local levels. When findings show AA perform better academically than other racial groups, there is no cause for concern, and funding is less likely to be allocated to focus on supporting AA students. As a result, certain AA subgroups miss out on needed opportunities that commonly target underrepresented and disadvantaged groups. A prominent example is inequitable opportunities in science, technology, engineering, and mathematics (STEM) fields for Southeast Asian students. Asians as a racial group are not considered an "underrepresented" group in STEM fields; thus, AA students are not targeted or eligible for opportunities to increase racial diversity in STEM.

STEM has become a priority in the U.S. to keep the economy competitive. Federal investments are poured into supporting STEM education efforts to build and diversify STEM fields. Such efforts most commonly target Black, Hispanic, and American Indian/Alaskan Native populations for being underrepresented groups. NHPI populations have received growing recognition as a disadvantaged minority group and are increasingly becoming included as an underrepresented minority group group. However, Asians are often excluded given national statistics that show they are overrepresented. In 2009, Asian workers represented 14% of the STEM workforce compared to 5% of the total workforce (Beede et al., 2011). The percentage of Asians in STEM fields has remained high between 2001 and 2014 in engineering (98%), computing (85%), and advanced manufacturing (83%) (Change the Equation, 2015). Asians are especially overrepresented in computer and mathematical occupations. In 2015, Asians represented 19% of those in computer and mathematical occupations, but only 9% of the professional workforce and 6% of the total workforce (Department of Professional Employees, 2015).

A closer look at STEM workforce data, however, indicates that the vast majority of Asian STEM workers (87%) are foreign-born (Beede et al., 2011). Among the 1.5 million foreign-born STEM workers in the U.S., 25% come from China and 11% from India. Many STEM positions are filled through the H-1B visa program established in the Immigration Act of 1990, where U.S. employers were allowed to hire foreign workers to fill positions requiring specialized skills and high educational attainment (American Immigration Council, 2018; Ruiz, 2017). The program currently has a cap of 65,000 visas a year, although the cap has been expanded or exceeded in past years. In addition, the first 20,000 applicants who earned a master's or doctorate's degree from a U.S. higher educational institution are exempt from the cap. Of the 430,000 initial H-1B guest worker visas (i.e., new visas opposed to continuing visas) issued between 2003 and 2011 in computer-related occupations, more than half (62%) of the visas went to workers from India and China (Department of Professional Employees, 2015). U.S. Census occupational data also show fewer Southeast Asians in Health and STEM fields than East Asians and South Asians (CARE, 2010). Less than 10% of Cambodian, Hmong, Lao, and Vietnamese workers are in health or STEM fields compared to over one-quarter of Asian Indian workers. While aggregated data show that Asians are indeed overrepresented in STEM, disaggregated data provide evidence that Southeast Asians are underrepresented.

Federal funding agencies, like the National Science Foundation (NSF), have yet to formally acknowledge Southeast Asian as an underrepresented group in STEM. In the 2017 report of statistics on underrepresented groups in science and engineering by the NSF and the National Center for Science and Engineering Statistics (NCSES), underrepresented groups were introduced as follows:

> The representation of certain groups of people in science and engineering (S&E) education and employment differs from their representation in the U.S. population. Women, persons with disabilities, and three racial and ethnic groups—Blacks, Hispanics, and American Indians or Alaska Natives—are underrepresented in science and engineering. (p. 2)

There is some evidence that exceptions are being made. With growing use of disaggregated data showing disparities within the Asian population, there are scholars who explicitly include Southeast Asian in their definition of underrepresented groups in STEM (Byars-Winston et al., 2010; Museus & Liverman, 2010; Museus, Palmer, Davis, & Maramba, 2011). Byars-Winston et al. (2010) states:

> Asian Americans in general have the highest U.S. college graduation rates, with a large percentage in STEM fields. However, disaggregated data reveal that Southeast Asians—largely Cambodian, Vietnamese, Hmong, and Lao individuals—have the highest high school drop-out rates and the lowest college graduation rates of all U.S. ethnic groups, on the basis of the 2006 American Community Survey from the U.S. Census Bureau. Further, Southeast Asians are least likely of all Asian Americans to be in managerial and professional occupations (where most STEM fields are groups). It is likely, on the basis of these statistics, Southeast Asians are underrepresented in science and engineering. (p. 205)

Access to higher education, in general, is an area where Southeast Asian students need support. Wright and Boun (2001) found that Southeast Asian college students often depended on others outside of their family (e.g., teachers, counselors, outreach programs) for information and help to get to college. With many Southeast Asian adults lacking English proficiency and higher education experiences, much less any formal education, Southeast Asian parents often do not have the skills or knowledge to help their children navigate higher education. Without the recognition of Southeast Asian being an "underrepresented" group—not just in STEM, but in higher education as well—they can be overlooked in college recruiting efforts and educational funding. More research illuminating the disparities that exist across Asian subgroups and greater advocacy is needed to ensure equitable opportunities for AA students.

Validity of inferences about Asian Americans and academic achievement. It is imperative to examine the validity of findings depicting AA as high achievers given the adverse impacts on AA students. Educational research is used to make policy decisions impacting access to and opportunities for students in the U.S. Much of the racial disparities research, especially on a national level, does not give a complete picture of which groups are falling behind and need targeted support.

For the examination of validity, I focus on the validity of inferences made from racial disparities research about AA rather than the technical aspects of measuring academic achievement as a construct. Much of the literature on validity is rooted in educational measurement and focuses on tests and test scores, however, I believe validity arguments can extend to other measures of academic achievement as well, like grade point averages; high school graduation and drop-out rates; and college enrollment, persistence, and graduation rates because these measures are commonly used as proxies of academic achievement. Thus, while I discuss validity in terms of test use and test scores based on relevant literature about validity—my arguments apply to inferences made from using all measures of academic achievement.

In educational measurement, test use and the consequences of test use play a large role in constructing the meaning of test scores. The current edition of the Standards for Educational and Psychological Testing defines validity as “the degree to which evidence and theory support the interpretations of test for proposed uses of tests” (American

Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME), 2014, p. 11). Messick (1989) defines validity as "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment" (p. 13). Thus, it is the uses and inferences made that need validation rather than the scores or measure itself.

Between the 1960s through 1980s, validity was popularized as a unitary concept focusing on construct-related validity, the extent a measure means what it is intended to mean (Kane, 2001; Messick, 1989; Thorndike & Thorndike-Christ, 2010). Prior to that, construct validity was not often perceived as a broader framework of validity, but as one of three forms of validity evidence that also included content-related evidence (the extent to which appropriate content is included to define a measure) and criterion-related evidence (the accuracy of a measure based on its relationship with other measures that should predict or discriminate the measure). In the unified framework of validity, construct validity encompasses content- and criterion-related evidence. Messick (1989), one of the most prominent validity theorists, stated, "Because content- and criterion-related evidence contribute to score meaning, they have come to be recognized as parts of construct validity. In a sense, then, this leaves only one category, namely, construct-related evidence" (p. 20). Kane (2001), another prominent validity theorist, identified four aspects of this validity framework, including the following:

1) The consideration of the plausibility of a proposed interpretation and uses of test scores (i.e., measures)

2) Proposed interpretations involve an extended analysis of inferences and assumptions, as well as a rationale for the proposed inferences in light of competing interpretations
3) The evaluation of the consequences of test use
4) Validity as an integrated evaluation of the interpretation (p. 328-329)

This framework is used to assess the validity of conclusions about AA students and consequences resulting from studies assessing academic achievement across racial groups.

The use of measures for assessing academic achievement is valid in the attempt to identify racial disparities in education. The different measures are plausible indicators of academic achievement in gauging subject competencies, coursework performance, and academic outcomes. The issue, however, lies in competing conclusions from studies using aggregated and disaggregated data. Studies using aggregated data show that AA students outperform other racial groups. The interpretation is believable given it is what the data show, but conclusions should be interpreted against other available evidence when assessing the validity of inferences (Kane, 2001). Three standards of educational and psychological testing (13.5, 13.6, 13.9) emphasize this need to prevent the misinterpretation of conclusions and appropriately inform decision-making (AERA, APA, & NCME, 2014). Standard 13.6 specifically states, "Reports of group differences in test performance should be accompanied by relevant contextual information, where possible, to enable meaningful interpretation of the differences. If appropriate contextual information is not available, users should be cautioned against misinterpretation" (p. 212). Educational researchers and program evaluators assessing racial disparities have the responsibility to assess the validity of their conclusions in light of the available research showing disparities within the AA population using disaggregated data.

Researchers and evaluators must also consider the consequences of the inferences they make—both intended and unintended (Kane, 2001; Messick, 1989). Within the unified framework of validity, consequences are often referred to decisions about individuals made from interpreting test scores, such as for admissions or placement purposes. However, I refer to consequences on a broader level having any kind of social impact, not just an impact at an individual level. Consequences as an aspect of validity itself have been highly debated in the field of educational measurement. The topic has been the focus of three journal issues:

1) *Educational Measurement*, 1997, 16(2)
2) *Educational Measurement*, 1998, 17(2)
3) *Assessment in Education: Principles, Policy & Practice*, 2016, 23(2)

Advocates argue that consequences are imperatively linked to construct validity and the process of validating test use (Kane, 2001; Kane, 2013; Linn, 1997; Messick, 1989; Shepard, 1997). The rationale to use a test for a particular purpose relies on the accuracy of score interpretation and the appropriateness of decisions made from such interpretations. Shepard (1997) points out that some test uses are to provide information for greater understanding, while some test uses have explicit intended consequences. She argues that when tests are used for decision-making, then consequences are inherent in the meaning of test scores and the framework of construct validity.

Opponents of consequences as an aspect of validity recognize the importance of evaluating consequences, but contend it does not need to be included in the process of test validation (Mehrens, 1997; Popham, 1997; Reckase, 1998). Popham (1997) perceived that assessing consequences would further cloud the concept of validity and confuse educational practitioners. He advocated for a simple definition of validity focused on the "accuracy of score-based inferences" (p. 10). He further believed consequences were the responsibility of test developers and users, where developers should examine and address potential consequences as part of distributing a test, while users had the responsibility of considering whether their use of a test was justified given the potential consequences.

Regardless of the different views, there is agreement that consequences of test use are important and should be addressed—especially in the presence of unintended consequences. Messick (1995) stated,

> Because performance assessments promise potential benefits for teaching and learning, it is important to accrue evidence of such positive consequences as well as evidence that adverse consequences are minimal. (p. 7)

To evaluate consequences, Kane (2001) suggests separating the interpretative argument into two parts: (1) a descriptive part focusing on the inferences made about individuals, and (2) the prescriptive part focusing on the decisions made from those inferences. Inferences made about AA students being high academic achievers have been heavily based on aggregated data that fail to provide a complete picture of AA and their academic experiences. The inferences themselves and the generalization of these inferences are further flawed based on growing evidence that show disparities across AA subgroups. The decisions that have been made from such inferences have led to adverse impacts for particular Asian subgroups, as discussed in the prior section. It is not just a matter of individual differences within the distribution of academic achievement, but distinct subgroups—like SEA—are not faring as well as other AA groups and even other racial groups. There exist inequitable opportunities for Southeast Asian students, despite the fact that they experience many of the same challenges and barriers as groups designated as underrepresented. Policy makers and stakeholders must have more complete information about the education of AA students to avoid increasing inequities.

The validity of inferences is susceptible to change in light of further research and new evidence (Messick, 1989). There is not enough validity evidence to continue to infer that all AA students are achieving high levels of academic success. Moreover, there is also not enough evidence to justify the lack of support and investment in the education of AA students. There is no question that aggregated data conceal disparities within the AA population. There is so much heterogeneity underneath research touting the academic success of AA over other racial groups. It has led to a lack of knowledge and understanding of the academic needs of AA students. In light of that, it is important for educational researchers, scholars, program evaluators, data scientists, and others to be equipped with appropriate knowledge and tools to move the current discourse on AA and academic achievement forward and provide a more valid and complete understanding of educational disparities.

Disaggregating Data on Asian Americans: How can it be done?

Aggregated data are flawed in accurately depicting academic achievement among AA students. For many years, scholars and advocates aware of disparities across AA subgroups have called for more nuanced data collection and analysis methods. There is growing research and advocacy around using disaggregated data for the AA population.

The White House Initiative on Asian Americans and Pacific Islanders (WHIAAPI) is an executive order first signed by former President Bill Clinton in 1999 and reestablished by former President Barack Obama in 2009 and President Trump in 2017. The initiative is active in promoting greater access to and participation in federal programming for Asians and Pacific Islanders (https://sites.ed.gov/aapi/). The National Commission on Asian American and Pacific Islander Research in Education (CARE) has also been active in engaging and informing the research community about the needs and challenges of Asian and Pacific Islander students (http://care.gseis.ucla.edu/). The organization is a strong advocate for the use of disaggregated data on AA and NHPI populations to inform education policy. Together, these two organizations launched a data quality campaign in 2013 called iCount (CARE, 2013). The iCount campaign seeks to:

1) Increase awareness of disparities across AA and NHPI subgroups
2) Provide models for collecting and reporting disaggregated data
3) Work collaboratively within the education field towards the implementation of practices to collect and report disaggregated data

This national effort was inspired by a smaller scale, but successful campaign for disaggregating data within the University of California (UC) college campuses. In 2006, the Asian Pacific Coalition, a group of 21 Asian American and Pacific Islander student organizations at the UC - Los Angeles (UCLA), initiated the Count Me In campaign (CARE, 2013). The campaign was in response to an article published in the UCLA student newspaper, *The Daily Bruin*, about AA students outnumbering White students for the first time (Doshi, 2006). Asian students were grouped in the article as Chinese, East Indian/Pakistani, Filipino, Japanese, Korean, Vietnamese, and Other Asian. The objective of the Count Me In campaign was to disaggregate the “Other Asian” category into 10 additional Asian subgroups to draw attention to the underrepresentation of Southeast Asian and NHPI students (CARE, 2013). As the campaign kicked off, student groups, faculty, and staff across UC campuses banded together. In the fall of 2007, UC administrators announced that the number of Asian sub-groups on the undergraduate application would expand from 8 categories to 23 categories—among the new categories were Cambodian and Lao. The changes were implemented in the 2009-10 school year. The success of the Count Me In Campaign was pivotal in the effort to advocate for the use of disaggregated data and shed light on existing disparities across AA subgroups. It serves as a model for how to track disaggregated data on AA subgroups, but also supports the fact that institutional change is possible to improve data systems that heavily impact educational policy.

There is growing political advocacy for the use of disaggregated data. As of December 2018, five states have passed legislation bills to disaggregate data on AA and NHPI, including California, Hawaii, Minnesota, Rhode Island, and Washington (Table 1). Advocates have celebrated and praised the bills as a step towards more nuanced data and greater knowledge to serve AA communities. Hawaii was the first state to pass legislation on disaggregating data on AA in 2012. The Hawaii Department of Education—which comprises only one school district—has collected disaggregated data for over 20 years (National Forum on Education Statistics, 2016). Data disaggregation bills vary across the five states. Minnesota and Washington are the only states to expand racial and ethnic subgroup disaggregation to other groups of color. Further, California is the only state to excludes educational agencies from collecting and reporting disaggregated data.

Table 1

Proposed or passed state legislation bills related to disaggregating data by racial and ethnic subgroup.

State	Details
California	The AHEAD (Accounting for Health and Education in Asian and Pacific Islander Demographics) Act or Assembly Bill 1726 passed in 2016 Selected state agencies, including the Department of Industrial Relations, the Department of Fair Employment and Housing, and the State Department of Public Health, will collect and use disaggregated data for Asian and Pacific Islander groups in demographic reporting. http://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201520160AB1726
Connecticut [a]	Senate Bill 465 proposed in 2018 Proposes state health agencies to collect and share demographic racial and ethnic subgroup data for Asian, African American, Hispanic (regardless of race), and Native Hawaiian and Pacific Islander individuals to reduce disparities in the health care system. https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&which_year=2018&bill_num=465
Hawaii	Senate Bill 2174 and House Bill 1983 passed in 2012 Requires state agencies, boards, or commissions to disaggregate data for: 1) Native Hawaiians 2) Each major Asian group, including Chinese, Japanese, Filipino, Korean, Vietnamese, Asian Indian, Lao, Cambodian, Bangladeshi, Hmong, Indonesian, Malaysian, Pakistani, Sri Lankan, Taiwanese, and Thai 3) Each major Pacific Islander group, including Samoan, Fijian, Tongan, and Guamanian. Disaggregated data is to be included in demographic reports and made public. https://www.capitol.hawaii.gov/Archives/measure_indiv_Archives.aspx?billtype=SB&billnumber=2174&year=2012 https://www.capitol.hawaii.gov/Archives/measure_indiv_Archives.aspx?billtype=HB&billnumber=1983&year=2012 https://www.capitol.hawaii.gov/session2012/bills/SB2174_.htm https://www.capitol.hawaii.gov/session2012/bills/HB1983_.HTM

[a] In 2018, Senate Bill 359 was also proposed in 2018 to prohibit the collection and reporting of racial and ethnic subgroup data by state education agencies unless required by federal law or collected for the entire student population.
(https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&which_year=2018&bill_num=359)

Table 1

Proposed or passed state legislation bills related to disaggregating data by racial and ethnic subgroup (continued…).

State	Details
Massachusetts	House Bill 3361 proposed in 2017 Proposes that state agencies, quasi-state agencies, entities created by state statute and sub-divisions of state agencies will use U.S. Census Bureau categories to identify Asian American and Pacific Islanders in data collection and reporting; with separate reporting for the five largest groups. https://malegislature.gov/Bills/190/H3361
Minnesota	All Kids Count Act passed 2016 and revised in 2017 Focuses on disaggregating education data for the seven largest Asian and Pacific Islander groups, three largest American Indian groups, seven largest Hispanic groups, and five largest Black and African American groups based on the most recent state demographer's report. Senate File 2597 in 2016 https://www.revisor.mn.gov/bills/bill.php?f=SF2597&y=2016&ssn=0&b=senate Senate File 1847 in 2017 https://www.revisor.mn.gov/bills/bill.php?f=SF1847&b=senate&y=2017&ssn=0
New York	Assembly Bill A7352 proposed in 2017 Proposes that state agencies, boards, or commissions that collect demographic data related to ancestry or ethnic origin begin to collect and report disaggregated data for major Asian and Pacific Islander groups. https://www.nysenate.gov/legislation/bills/2017/a7352/amendment/original
Rhode Island	All Students Count Act or House Bill 5453 passed in 2017 Requires the Department of Elementary and Secondary Education to use disaggregated data for the Asian population in data collection and reporting. Selected groups may include, but should not be limited to: Cambodian, Filipino, Hmong, Lao, Vietnamese, and Other Southeast Asian groups. This requirement only applies to students attending public schools. http://webserver.rilin.state.ri.us/billtext17/housetext17/h5453.htm

Table 1
Proposed or passed state legislation bills related to disaggregating data by racial and ethnic subgroup (continued…).

State	Details
Washington	House Bill 1541 passed in 2016 Requires public school districts and education data agencies to collect and report specified subgroup data for Black, Asian, White, and multiracial students. http://app.leg.wa.gov/RCW/default.aspx?cite=28A.300.042 http://lawfilesext.leg.wa.gov/biennium/2015-16/Pdf/Bills/Session%20Laws/House/1541-S4.SL.pdf

Efforts to pass similar data disaggregation legislation are ongoing in Connecticut, Massachusetts, and New York. Interestingly, Connecticut had two bills proposed in 2018 related to data disaggregation: one to disaggregate data for state health agencies and the other to prohibit data disaggregation for state education agencies (Kara, 2018; Moser, 2018). There has been a strong opposition to data disaggregation bills from the Chinese American community. Media news articles covering protests of data disaggregation bills report Chinese American parents have strong concerns about how the data will be used (Fuchs, 2016; Fuchs, 2017a; Fuchs, 2017b; Fuchs, 2018; Kara, 2018; Li, 2018; Moser, 2018; Wang, 2016b; Wang, 2017). Their opposition is based on three key issues:

1) Asians are being singled out to create an "Asian registry" that could be used for discriminatory policies and actions, such as what happened with the Chinese Exclusion Act of 1882 and the internment of Japanese Americans during World War II.
2) There is potential for data disaggregation to advance race-based college admissions policies or undermine current legislation prohibiting such policies,

which would put high-achieving Chinese American students at a disadvantage for enrollment at competitive higher education institutions.

3) Data disaggregation can lead to re-distributing federal funding where Southeast Asian subgroups would benefit, while others would not. Opponents generally believe investing more federal assistance into communities will not improve educational outcomes.

Policy makers face a difficult dilemma and will need to be prepared to address concerns and fears among opponents, while at the same time meet the data needs of advocates who do not see their constituents represented in aggregated data.

To some extent, there is recognition of the need for disaggregated data by the U.S. Department of Education (ED). In November 2016, ED awarded $836,000 to three state grantees (Minnesota, Washington, and Hawaii) to improve data collection systems and disaggregate data on Asian and Pacific Islander students; the project is known as the Asian American Pacific Islander Data Disaggregation Initiative (U.S. Department of Education, 2016). When the initiative was initially announced, former Secretary of Education John King said (Wang, 2016a, video clip):

> We know that many Asian Americans and Pacific Islanders (AAPI) face the model minority myth, the notion that virtually all AAPI have access to a quality education and are affluent—which has prevented AAPI communities from fully benefitting from federal programs and resources that support vulnerable and underserved people. In reality, the AAPI community is not a monolithic group and AAPI face unique challenges, including in education.

Several years prior to establishing the initiative, the ED put out a request for information about practices and policies of collecting data on AA and NHPI students in 2012 (Dinh, 2013). Comments were received from many individuals (e.g., students and parents), educational agencies (e.g., school districts and state departments of education), and advocacy groups. Responses emphasized a growing need for disaggregated data to better serve and support AA students, especially in geographical areas where the AA community was diverse (e.g., California and Minnesota). There were many challenges and concerns mentioned, including small numbers, the expense of infrastructure changes (e.g., staff training, revising student enrollment forms, and revising databases), the scope of disaggregation to other racial groups, and inconsistencies in data collection across educational institutions.

Collecting and using disaggregated data is certainly not without challenges. There are many ways disaggregated data are collected (Dinh, 2013; Islam et al., 2010). Commonly, birthplace and/or home language are collected as proxies for AA ethnic subgroups. Such proxies, however, are becoming less accurate with the growth of AA children born in the U.S. and the loss of ethnic languages in preference of English. Moreover, birthplace is not equivalent to ethnic subgroup; for example, Chinese immigrants born in Vietnam, Asian Indian immigrants born in Malaysia, or Hmong refugees born in Laos or Thailand. Rather than use proxies, it is more accurate to collect specific AA subgroup information. For example, as mentioned previously, the 2010 U.S. Census collected data on 25 different AA subgroups and 20 NHPI subgroups (Hixson, et al., 2012; Hoeffel et al., 2012). Additionally, UC campuses collect student data using 23

AA and NHPI subgroups (CARE, 2013). Yet doing so results in many subgroups for consideration, and there are challenges to using such nuanced data.

A common issue with using disaggregated data on AA is small sample sizes (Islam et al., 2010; Okazaki & Sue, 1995). Small sample sizes can deter the collection of disaggregated data altogether, but also pose a threat in identifying individuals (Dinh, 2013). These are constant dilemmas for educational agencies and institutions, as well as for program evaluation as it is common for programs to serve a small number of participants, e.g., fewer than 30. For research design, small samples pose a threat to statistical conclusion validity, the appropriateness of statistical methods to conclude that a relationship exists among variables (Shadish, Cook, & Campbell, 2002). When samples sizes are too small, it decreases statistical power. Shadish, Cook, and Campbell (2002) define statistical power as "the probability of finding an effect when an effect exists" (p. 510). For statistical significance testing, small samples may lead researchers and evaluators to conclude no effects or differences where one exists (known as type II error). When working with disaggregated data, researchers and program evaluators must decide whether small groups will be excluded or combined with other groups (Okazaki & Sue, 1995). Combining AA subgroups can make analyses simpler by comparing fewer groups, but it also makes groups more heterogeneous, introducing more variation in analyses.

There are ways to remedy small sample sizes. Oversampling, known as unequal probability sampling, is frequently used in survey research designs as a way to increase the representativeness of groups of interest, target groups who are difficult to reach, and increase statistical power for small groups (Dillman, Smyth, & Christian, 2009; Holland & Palaniappan, 2012; Islam et al., 2010; Kalton, 2009; Marpsat & Razafindratsmia, 2010). In a review of the sampling designs used in international and national education datasets sponsored by the National Center for Education Statistics (NCES), there were 4 of 13 data sets that oversampled AA students: the Early Childhood Longitudinal Survey (ECLS), Educational Longitudinal Study of 2002 (ELS), the High School Longitudinal Study of 2009 (HSLS), and the National Educational Longitudinal Study (NELS) of 1988 (Table 2). These four studies also collected disaggregated data on AA students, an indication that there is some recognition of educational disparities within the AA community in national education research efforts. Unfortunately there was no information provided about whether Asians were oversampled based on subgroup. Results and findings can be sensitive to which Asian subgroups were oversampled given that East and South Asians typically have higher academic achievement than Southeast Asians. One can assume the oversampling was done proportionally based on the make-up of Asian subgroups in the U.S. since the data are nationally representative, but without more details it is impossible to know for sure.

Another method to increase sample sizes is pooling or linking data (Islam et al., 2010). Doing so essentially combines student-level data across years, cohorts, or studies for subgroups to have appropriate numbers for statistical power. Pooling national datasets across years has been used in public health research to identify the prevalence of health issues and disparities in health access across AA subgroups (Islam et al., 2010). As a result, public health efforts and support can be more targeted and tailored to specific AA subpopulations. In educational research, pooling techniques work well with assessment data collection that is periodically administered over time to different cohorts. In their analysis of disaggregated AA subgroups, Pang et al. (2001) pooled together California

Table 2

Sampling and data collection methods for Asian American populations across national and international educational data sets.

Name of data set	Timeframe & cohorts	Oversamples Asian students	Collects data detailing Asian subgroup	Collects proxy information to identify Asian subgroup
Baccalaureate and Beyond Longitudinal Study (B&B)	2008-2012	No	No	No
Beginning Postsecondary Students (BPS)	Several cohorts: 1990-1994; 1996-2001; 2004-2009; 2012-2017	No	No	No
Early Childhood Longitudinal Survey (ECLS)	Several cohorts: Birth 2001-2006, Kindergarten 1998-1990, Kindergarten 2010-2011	Yes	Yes: Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Other Asian (Specify)	Yes: Home language
Educational Longitudinal Study of 2002 (ELS)	2002-2012	Yes	Yes: Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Other Asian (Specify)	Yes: Native language
High School and Beyond (HS&B)	1980-1993	No	No	No
High School Longitudinal Study of 2009 (HSLS)	2009-2016	Yes	Yes: Chinese, Filipino, Southeast Asian, South Asian, Other Asian	Yes: Native language
National Assessment of Educational Progress (NAEP)	1969-2024	No	No	No

Source: Information retrieved from respective studies' websites and technical manuals; see https://nces.ed.gov/surveys/.

Table 2
Sampling and data collection methods for Asian American populations across national and international educational data sets (continued…).

Name of data set	Timeframe & cohorts	Oversamples Asian students	Collects data detailing Asian subgroup	Collects proxy information to identify Asian subgroup
National Educational Longitudinal Study (NELS)	1988-2000	Yes	Yes: Chinese, Filipino, Japanese, Korean, Southeast Asian, Pacific Islander, South Asian, West Asian, Middle Eastern, Other Asian	Yes: Native language
National Household Education Survey (NHES)	Data collection completed between 1991-2007, but varies by survey topic	No	No	Yes: Birth country and native language (varies by topic)
National Longitudinal Study (NLS)	1972-1986	No	No	No
National Postsecondary Student Aid Study (NPSAS)	Every three years between 1987-1996 Every four years between 1996-2012 Every two years starting in 2016	No	No	No
Program for International Student Assessment (PISA)	2000 and every 3 years after	No	No	Yes: Birth country & home language
Trends in International Mathematics and Science Study (TIMSS)	1995 and every 4 years	No	No	Yes: Home language

Source: Information retrieved from respective studies' websites and technical manuals; see https://nces.ed.gov/surveys/.

standardized assessment data across several years, 2003 to 2008. Sample sizes across AA subgroups ranged from 6,763 for Lao students to 63,860 for Filipino students. With larger sample sizes, Pang et al. (2011) used statistical significance testing methods to assess group differences, first using an independent Welch-Satterthwaite t-test to examine differences between the aggregated AA and Pacific Islander group compared to Whites, and then using analysis of variance (ANOVA) to examine differences across AA and Pacific Islander subgroups compared to Whites. Without different cohorts, however, pooling is not an option. Many national educational datasets are longitudinal with only one cohort. Linking longitudinal datasets may be possible, however, it would only be relevant if AA subgroup information is collected.

Advocacy efforts for using disaggregated data continue to grow to provide AA students with equitable support and opportunities. There are many opportunities in educational research and evaluation to collect and use disaggregated data for the AA population, but there is also a need for greater consistency and guidance around doing so.

Chapter 3: Methods

This study examines the use of aggregated and disaggregated groups for Asian American (AA) students while assessing disparities in math achievement. Similar analyses are completed using a national-level data set, the High School Longitudinal Study of 2009 (HSLS), and a state-level data set, the Minnesota Comprehensive Assessments-II (MCA-II). This comparison will help to assess the presence of math achievement disparities across Asian subgroups and the need to use disaggregated data.

The analysis for both data sets focus on the cohort of 9th graders in the 2009-10 school year. Two criteria were used to select comparable student samples from both data sets, including:

1) Students attended a public school during assessment periods
2) Students did not fail or skip a grade between assessment periods

The primary group of interest in this study is AA students. For comparison, two additional racial groups are included in the analyses: White students, and Black students. These two groups were intentionally chosen because White students are often used as the comparison group when assessing racial disparities, while Black students typically experience the largest achievement gap from White students and are considered an underrepresented minority group. Students who identified as having Hispanic and/or Latino ethnicity were excluded from the sample. I acknowledge that there are challenges in data use and issues of inequity for other race and ethnicity groups (American Indian/Alaskan Native, NHPI, individuals of mixed race, and individuals of Hispanic/Latino descent). However, for the scope of this study, the number of race and ethnicity groups used for comparison was reduced to focus on disaggregating data for AA students.

The following outlines the research questions and hypotheses for this study:

1) To what extent are the distributions of math scores for Asian, White, and Black student groups bimodal/multimodal? If there is evidence of multimodality, how do the mode groups differ across demographics?

    Hypothesis: The distribution of math scores will be bimodal/multimodal for Asian students and unimodal for White students and Black students.

2) To what extent are there differences in mean math scores across AA subgroups?

    Hypothesis: Southeast Asian students will have lower mean math scores than East Asian students and South Asian students.

3) To what extent are there differences in math scores during high school by race while grouping Asians as an aggregated group compared to disaggregating Asians into subgroups?

    Hypothesis: Use of an aggregated Asian group will result in Asians having the highest math achievement compared to White students and Black students. However, using disaggregated Asian subgroups will show varying levels of math achievement—where East Asian students and South Asian students will perform higher than White students and Black students, while Southeast Asian students will perform lower than White students, but similar to Black students.

The distribution of math scores will first be examined. Researchers advocating for greater use of disaggregated data have cited a bimodal distribution of academic attainment and achievement for Asians, where subgroups like Chinese and Asian Indian have higher educational attainment and achievement levels compared to Southeast Asian subgroups (Hartlep, Morgan, & Hodge, 2015; Hune & Chan, 1997; Teranishi, 2010). Factors like sex, immigrant generational status, first-generation college student status, English and native language proficiency, and U.S. region also contribute to educational disparities across Asian subgroups (Galindo & Pong, 2014; Hall, 2009; Hartlep, Morgan, & Hodge, 2015; Hune & Chan, 1997; Kao, 1995; Lee & Zhou, 2015; Ng, Lee, & Pak, 2007; Pak, Maramba, & Hernandez, 2014; Teranishi, 2010). The presence of a bimodal distribution among AA students is problematic because it is an indication that the sample is not homogeneous enough for comparison to other groups. Further, if the modes largely consist of specific Asian subgroups, it warrants the use of disaggregated data to more accurately capture disparities. Histograms were first run to examine the distribution of math scores using IBM SPSS Version 24. The shape and number of modes present in histograms are sensitive to the bandwidth used, i.e., the interval in which frequencies are grouped for display. The automatic bandwidth computed by SPSS was used to descriptively examine the distribution of scores for each racial group. Detailed information about how the automatic bandwidth is computed was not available in the SPSS online manual, but in an IBM support community it states, “The number of bars is calculated by an algorithm that uses statistical theory to suggest a number of bars that is optimal for a data set of the size provided, under an assumption of normally-distributed values” (http://www-01.ibm.com/support/docview.wss?uid=swg21480583).

The Silverman Test of Multimodality (Silverman, 1981) was then used to estimate the number of modes present in the distributions of math scores for each racial group. This analysis was completed using the "silvermantest" package with R version 3.3.3 (see https://www.mathematik.unimarburg.de/~stochastik/R_packages/silvermantest_manual.pdf). The Silverman Test identifies the number of modes by estimating the probability density function using kernel density estimates (Adereth, 2014; Hall & York, 2001; Silverman, 1981). The null hypothesis of the Silverman Test is that the probability density function has at most a specified number of modes, k, while the alternative hypothesis is the probability density function has more than k modes. In this sense, the mode is the value with the highest probability of being observed (Adereth, 2014). Bootstrapping is used to sample values with replacement and test whether there are at most k modes. The kernel density estimates, i.e., the estimated normal density function centered at each value, is computed and smoothed to estimate the underlying probability density function for each replication. The resulting p-value for the Silverman Test is the fraction of the number of times the null hypothesis is rejected over the total number of replications in the simulation (Hall & York, 2001). For any given k modes, a minimum critical bandwidth, h, can be computed from the smoothed kernel density function, where h is the critical bandwidth at which a new mode appears (Adereth, 2014; Hall & York, 2001; Silverman, 1981). The distribution of math scores for each racial group at two time points was tested for up to two modes using the default of 999 bootstrap replications. When $k = 1$ was tested, an adjustment was used to improve the accuracy of the p-value (Hall & York, 2001). Crosstabs were run using IBM SPSS Statistics Version 24 to examine whether mode groups differ across demographics using Chi-square Tests of

Independence or the Fisher's Exact Test for 2x2 crosstabs. Z-tests for column proportions with Bonferroni adjustments to the p-value will also be run to identify significant differences across mode groups.

For the second research question, boxplots and an analysis of variance (ANOVA) were completed to assess differences in math achievement across Asian subgroups. Different classification methods for Asian subgroups were used in the ANOVA, when possible, to guide how to disaggregate Asian subgroups. The ANOVA was completed using IBM SPSS Statistics Version 24.

Linear mixed effects (LME) modeling with maximum likelihood estimation was used to examine racial disparities in math achievement while comparing the use of aggregated versus disaggregated groups for AA students. Math achievement was assessed at two time points: 9^{th} and 11^{th} grade with the HSLS data, and 8^{th} and 11^{th} grade for the MCA data. The use of LME modeling accounts for the correlation between repeated measures (Fitzmaurice, Laird, & Ware, 2011). The random intercepts model was used to account for differences in the first math achievement score across subjects, while the slopes are fixed. Preliminary analyses supported the use of linear mixed modeling over fixed effects modeling to assess changes between the two grade levels, where the random intercept models had lower AIC and BIC fit statistic values than the fixed effects models. The analysis was completed using IBM SPSS Statistics Version 24 with the MIXED feature available in the advanced statistics module. The LME modeling used the default scaled identity covariance structure, which is the only error structure option for models with two time points. This covariance structure assumes constant variance and no correlation across measures (see https://www.ibm.com/support/knowledgecenter/en/SSLVMB_23.0.0/spss/advanced/covariance_structures.html). For each model, covariates were initially included and the model was reduced until all covariates significantly contributed to the model at an alpha level of $\alpha = 0.010$ for the HSLS data and $\alpha = 0.050$ for the MCA-II data. A more conservative alpha level is used for the HSLS data because weighting was applied to account for the complex sampling design. Details are explained in the following section.

The High School Longitudinal Study of 2009 (HSLS)

The HSLS of 2009 restricted use data set was used for this study. It is part of a series of longitudinal studies focused on secondary education from the National Center for Education Statistics (NCES). The study explores student trajectories from high school to post-secondary education (and eventually trajectories into the workforce) with an emphasis on STEM pathways (Ingels et al., 2011). It is one of the few and the most current NCES data sets to collect disaggregated data for Asian students—making it relevant and appropriate for use in this study. The HSLS began in the 2009-10 school year with a nationally representative sample of fall-enrolled 9th graders in the U.S. The study included questionnaires completed by students, parents, teachers (both math and science), school administrators, and counselors.

Sampling design. The HSLS used a stratified, two-stage random sampling design (Ingels et al., 2011). In the first stage, schools were stratified by type (public—including charters, private—Catholic, and private—other), U.S. region (Northeast, Midwest, South, and West), and locale (city, suburban, town, and rural). Probability proportional to size sampling was used to select schools randomly from each stratum to be representative of the 50 states and the District of Columbia. Private Catholic schools were oversampled.

Further, schools needed to meet some eligibility criteria to be part of the study, such as having both the 9th and 11th grades (see Ingels et al., 2011 for complete list of eligibility criteria). A total of 1,889 eligible schools were randomly selected for the study, of which 944 schools participated.

In the second stage, 9th grade students enrolled in the fall of 2009 were sampled within participating schools. Students were first stratified by race (Hispanic, Asian, Black, and Other) then randomly selected for the study. Asian students were oversampled to yield a sufficient sample size for analysis. A total of 24,658 students met study eligibility criteria and were capable of completing the questionnaire and math assessment. Of those students, 21,444 completed the base-year questionnaire and 20,781 completed the math assessment.

Missing data. The HSLS data set includes several codes for missing data: -7 for legitimate skips, -8 for not completing the questionnaire or assessment, and -9 for missing data (Ingels et al., 2011). The HSLS base-year file documentation states that there were not "high levels" of missing data, but acknowledges that missing data were unlikely to be missing completely at random (i.e., unrelated to any variable or characteristic) (Ingels et al., 2011, p. 162). Item imputation procedures replacing missing data were used to allow for complete-case analysis. Single-value imputation was used to replace missing data for 18 demographic variables (some of which are used in this study—e.g., race and parents' highest level of education), while a model-based multiple imputation procedure was used to impute data for math theta scores, the standard error of measurement for math theta scores, and socioeconomic status (SES). More details about item imputation can be found in the HSLS: 09 base-year file documentation (Ingels et al., 2011).

Sample weight and design effect adjustment. The student level weight, W1W2STUDENT (w_i), was used to be representative of the population of 9th graders in the U.S. when analyzing the base year and first year follow up data. This study focuses on student level effects and does not perform multi-level modeling that would examine school level effects and account for the complex sampling design used in HSLS (e.g., stratification, clustering, and oversampling techniques). Researchers urge against treating study samples collected from complex sampling methods as a simple random sample (Osborn, 2011; Hahs-Vaughn, 2005; Hahs-Vaughn, 2006; Thomas & Heck, 2001). Complex sampling designs produce samples of students who are more similar to each other than if a simple random sample were drawn directly from the population. If the complex sampling design is not accounted for in examining student level effects, there is a risk of underestimating standard errors, violating the assumption of independent observations, and making Type I errors in significance testing (i.e., concluding a statistically significant effect or difference when it is not present).

AM software (see http://am.air.org/) is available for analyzing student-level effects using large-scale data derived from complex sampling strategies (Osborn, 2011; Hahs-Vaughn, 2005; Thomas & Heck, 2001). AM is a free statistical software from the American Institutes for Research (AIR) designed to account for complex sampling strategies by allowing researchers to identify the cluster, strata, and weight variables. Since the software is currently being tested and only available in a Beta version, adjustments to the student weight were made instead to account for the complex sampling design of the HSLS. Weighting adjustments are another way to account for bias in sampling design. When using this method, it is recommended to use a more conservative alpha level, $\alpha = 0.010$, for statistical hypothesis testing (Thomas & Heck, 2001). The student weight was first normalized to account for the original sample size:

$$Normalized\ weight = w_N = \frac{w_i}{\overline{w}} \tag{1}$$

This step is necessary because applying the student-level weight as is represents the population of 9th grade students in the U.S. Hypothesis testing using estimated large population sizes will likely result in statistically significant findings (Osborn, 2011; Hahs-Vaughn, 2005; Thomas & Heck, 2001). Thus, it is important to scale the weight and run the analysis on the original sample size instead. The weight was then divided by the design effect (DEFF) to account for the complex sampling method used in the HSLS:

$$Design\ effect\ adjusted\ weight = w_{DEFF} = \frac{w_N}{DEFF} \tag{2}$$

The DEFF is the ratio of estimated variances of a characteristic for a complex sample design to that of a simple random sample design of the same size (Hahs-Vaughn, 2006; Ingels et al., 2013). The average DEFF of 4.4 for analysis of all students from the HSLS first follow up manual was used (Ingels et al., 2013, p. 126).

Study sample. After the weight adjustments and sample selection criteria, the effective sample size included 170 Asian students, 610 Black/African American students, and 2,330 White students in the sample. Additional descriptive data are provided in Tables 3 and 4. Sample sizes are rounded to the nearest tenth to protect the privacy of the restricted-use HSLS data.

Table 3

Frequencies of student demographics from the HSLS data set.

Variable	Frequency (n)	Percent (%)
Race		
Asian	170	5.6
Black	610	19.7
White	2,330	74.8
Asian subgroup		
Chinese	40	23.0
Filipino	20	13.7
Southeast Asian	50	30.3
South Asian	40	20.9
Other Asian	20	12.1
Sex		
Male	1,550	49.6
Female	1,580	50.4
Immigrant generation		
1st generation	30	3.0
2nd generation	110	12.8
3rd generation or later	750	84.2
Dual Language – first language learned		
English only	2,890	92.9
Non-English language only	130	4.3
Both English & non-English equally	90	2.9
First-generation college student status		
Yes	670	25.8
No	1,920	74.2
Geographical region of the U.S.		
Northeast	590	19.0
Midwest	840	26.8
South	1,200	38.3
West	500	15.9
185% U.S. poverty threshold		
At or above 185% of poverty threshold	2,040	65.5
Below 185% of poverty threshold	1,070	34.5
Most advanced math course taken		
Pre-algebra or Algebra 1, 1A or 1B	100	4.1
Algebra II or III	1,100	40.8
Geometry or Analytic Geometry	310	11.5
Trigonometry	210	7.9
Pre-calculus or Analysis and Functions	550	20.6
AP Calculus AB or BC or Other Calculus	120	4.4
AP Statistics or Other Statistics or Probability	110	4.1
Integrated Math I-III or above	80	2.9
IB math standard or higher level	20	0.6
Other math course	90	3.3

Note. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

Table 3
Frequencies of student demographics from the HSLS data set (continued…).

Ever taken an advanced math course		
Yes	490	84.4
No	2,630	15.6

Note. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

Table 4
Descriptive statistics of HSLS math theta scores by grade.

Statistic	9th Grade	11th Grade
N	3,080	3,120
Minimum	24.1	25.3
Maximum	82.2	84.9
Mean	51.1	51.1
Median	50.9	50.9
Standard Deviation	9.9	9.9
Variation	98.6	98.7
Range	58.1	59.6
Kurtosis		
Statistic	-0.2	-0.3
Standard Error	0.1	0.1
Skewness		
Statistic	-0.1	0.1
Standard Error	0.0	0.0

Note. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

Variables. The following four demographic variables were self-reported by students: race, Asian subgroup, sex, and dual language. Race is a categorical, nominal variable where 1 = Asian, 2 = Black, and 3 = White. Asian subgroup is a categorical, nominal variable, where 1 = Chinese, 2 = Filipino, 3 = Southeast Asian, 4 = South Asian, and 5 = other Asian (e.g., Korean, Japanese and others). Asian subgroups were examined in the ANOVA using two grouping classifications: 1) the original five categories as provided by the HSLS, and 2) three categories consisting of East Asian, Southeast Asian, and South Asian. The Asian subgroup variable was condensed to include three categories because a theoretical reason for separating Filipino students from Southeast Asians was not provided in the HSLS data manuals. Since the Philippines is geographically located in Southeast Asia, Filipino students and Southeast Asian students were grouped together. To remain consistent with Southeast Asians and South Asians being categorized by geographic region, Chinese students and other Asian students were combined into a group as most students identifying as "other Asian" (76%) had one or both parents born in Korea or Japan. If ANOVA results showed that there are not significant differences in mean math scores between Filipino and Southeast Asian students and between Chinese students and other Asian students, the 3-category Asian subgroup rather than the 5-category Asian subgroup variable would be used in the LME modeling.

Sex is a dichotomous, nominal variable, where 1 = male and 2 = female. The dual language indicator is the first language students learned to speak. It is a categorical, nominal variable with 1 = English only, 2 = non-English language only, and 3 = both English and non-English language equally. The HSLS does not have a variable indicating a students' proficiency in English and/or some Other language. Immigrant generation and first-generation college student status were recoded from information provided by a parent in the parent questionnaire. Immigrant generation is a categorical, nominal variable, where 1 = first-generation immigrant, 2 = second-generation immigrant, and 3 = third-generation immigrant or later. First-generation immigrants include students who are foreign-born (i.e., born outside of the U.S., Puerto Rico, or other U.S. territory) and have two foreign-born parents. Second-generation immigrants are defined as native-born students (i.e., born in the U.S., including Puerto Rico and other U.S. territories) with at least one foreign-born parent. Third-generation immigrants or later include native-born students with two native-born parents. There were a small number of students who were foreign-born, but had one or both parents being native-born. For these students, those who had two native-born parents were coded as 3rd generation, while those with only one native-born parent were coded as 2nd generation. This definition aligns with U.S. Census Bureau's definition of immigrant generational status (see https://www.census.gov/topics/population/foreign-born/about.html). It was used for simplicity and to avoid having small sample sizes across many generational cohorts, but it is acknowledged that there are varying levels of acculturation within the first and second immigrant generations that influence educational experiences (see Rumbaut, 2004).

First-generation college student status is a dichotomous, nominal variable, where 0 = no and 1 = yes. A first-generation college student is defined as a student with neither parent nor guardian having a post-secondary education degree; in other words, both parents have a high school diploma/GED or less (Redford & Hoyer, 2017). If the highest level of education attained for both parents were a high school diploma/GED or less, the student received a code of 1 for being a first-generation college student. If one or both parents had at least an associate's degree, the student was coded as 0 for not being a first-generation college student.

Poverty threshold and geographic region were provided in the HSLS data set. This study uses the 185% U.S. Census poverty threshold rather than SES because the SES indicators do not account for household family size. AA households tend to be multi-generational, have more individuals working, and have larger family sizes (Fry & Passel, 2014; Pew Research Center, 2013; Ramakrishnan & Ahmad, 2014). Poverty thresholds are used to set poverty guidelines for eligibility in federal programs, like the National

School Lunch program (see https://www.census.gov/topics/income-poverty/poverty/about.html and https://aspe.hhs.gov/frequently-asked-questions-related-poverty-guidelines-and-poverty). The 185% poverty threshold is a dichotomous, nominal variable where 0 = at or above 185% of the poverty threshold and 1 = below 185% of the poverty threshold. Geographical region of the U.S. is the location of the school that students were sampled. It is a categorical, nominal variable, where 1 = Northeast, 2 = Midwest, 3 = South, and 4 = West.

Math coursework and advanced math course taking were self-reported by students in the student questionnaire in 11th grade, while math achievement score was provided by the HSLS data set from the 11th grade math assessment. In the questionnaire, students were given a list of math courses and asked to identify which ones they had taken. The data were used to identify the most advanced math course the student had taken and grouped accordingly from least to most advanced, where 1 = pre-algebra, 2 = algebra I, 2 = algebra II or III, 3 = geometry, 4 = trigonometry, 5 = pre-calculus or analysis and functions, 6 = calculus, 7 = statistics, 8 = integrated math I-III or above, 9 = International Baccalaureate (IB) math standard or higher level, and 8 = other math course. For the analysis, pre-algebra and algebra I were grouped based on the small number of students who took these courses. Integrated math and IB math courses were also grouped with the other math course category based on sample size. Other math courses included: math III, mathematical modeling and analysis, computer science, contemporary mathematics in context II or III, computer math, economics, math support courses, quantitative reasoning, and other miscellaneous courses. Advanced math course taking is a combination of aggregating three questions where students were asked if they had ever taken an Advanced Placement (AP), IB, and/or other courses for college credit in the subject math. The variable is coded as 0 = no and 1 = yes. Math achievement is the sole outcome variable for this study. The math theta score is a norm-referenced measurement of achievement measuring algebraic reasoning (Ingels et al., 2011, Ingels et al., 2013). It is suitable for comparing students' level of achievement relative to the student population. Students were assessed in 9th grade and 11th grade. The scores are a transformation of the math theta estimate to have a mean of 50 and a standard deviation of 10.

Statistical analysis and models. The following is an outline of the method, models, and statistical tests used to examine each research question. An alpha level of α = 0.010 was used for all statistical hypothesis testing.

1) To what extent are the distribution of math scores for Asian, White, and Black student groups bimodal/multimodal? If there is evidence of multimodality, how to do the mode groups differ across demographics?

Method:

1) Run Silverman's Test of Multimodality and examine density plots of the distribution of math scores.
2) Examine mode groups by completing crosstabs with chi-square tests and z-tests

Statistical tests:

Silverman's Test for Multimodality

H_o: $k \leq x$
H_1: $k > x$
where k = number of modes and x = 1, 2, or 3

Chi-square Test of Independence or Fisher's Exact Test for 2x2 crosstabs

H_o: Mode group is independent of demographic variable
H_1: Mode group is not independent of demographic variable

Z-test for column proportions adjusting for p-values using the Bonferroni method

Note: Since the HSLS data were weighted to account for the complex sampling design, new data frames were created with cases replicated by their weight multiplied by a factor of 10.

2) To what extent are there differences in mean math scores across AA subgroups?

Method: Run a one-way analysis of variance (ANOVA) with post-hoc tests comparing mean math scores across AA subgroups.

Model specification: $Y_{ij} = \mu_i + e_i$

where:

Y_{ij}= predicted math score for the *j*th student in the *i*th Asian subgroup

μ_i = mean math score for the *i*th Asian subgroup

e_i = error term for the *i*th Asian subgroup

Statistical tests: ANOVA F-test for:

H_0: μ_i are equal

H_1: two or more μ_i are not equal

Tukey or Games-Howell post-hoc tests depending on whether data meet homogeneity of variance assumption

3) To what extent are there differences in math scores during high school by race while grouping Asians an aggregated group compared to disaggregating Asians into multiple subgroups?

Method: Use LME modeling to run a random intercept model to examine changes in math achievement between 9th and 11th with the following variables:

- Race with an aggregated Asian group or disaggregated Asian subgroups (separate models were run with White students and Black students as the reference group for comparison with Asian subgroups)
- Sex (reference group = males)
- Immigrant generation (reference group = third-generation or later)
- Dual language (reference group = English only)

- First-generation college student status (reference group = No)
- U.S. geographic region (reference group = South)
- 185% poverty threshold (reference group = at or above the threshold)
- Most advanced math course taken (reference group = Algebra II/III)
- Advanced math course taking (reference group = yes)

Model specification: Level 1: Within-subjects

$$Y_{ij} = \beta_{0j} + \beta_1(Grade_{ij}) + \beta_n(X_{nij}) + e_{ij}$$

where:

i =	individual
j =	time
Y_{ij} =	predicted math score for the ith individual at time j
β_{0j} =	mean math score for individuals at the starting grade level, 9th grade
β_1 =	rate of change for the ith individual, i.e., slope
β_n =	beta coefficient for nth covariate
X_{nij} =	nth covariate for the ith individual at time j
e_{ij} =	error term for the ith individual at time j

Level 2: Between-subjects

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

where:

γ_{00} =	mean math score for the whole sample, i.e., grand mean
μ_{0j} =	error for random intercepts for time j

Full Model:

$$Y_{ij} = (\gamma_{00} + \mu_{0j}) + \beta_{1j}(Grade_{ij}) + \beta_{2j}(X_{2ij}) + \cdots + \beta_{nj}(X_{nij}) + e_{ij}$$

Minnesota Comprehensive Assessments (MCA) – II

MN was selected as a statewide comparison for this study because it is home to a diverse Asian community with a large Southeast Asian population. Asians make up 5% of MN's population overall (U.S. Census Bureau, 2017 Table B02001), with Southeast Asians comprising over half of the Asian population (58%) (U.S. Census Bureau, 2017 Table B02015). Furthermore, as previously mentioned, MN is a grantee of the Asian American Pacific Islander Data Disaggregation Initiative sponsored by the U.S. Department of Education. As part of this initiative, the MN Department of Education (MDE) is embarking on developing and improving its data collection systems to disaggregate data for not only Asian and Pacific Islanders students, but for students from other races and ethnicities as well (For more information, see https://education.mn.gov/MDE/dse/count/).

The MCA-II[2] is a set of the statewide accountability tests used in MN. The MCA-II mathematics assessment was used to measure math achievement and progress towards meeting math standards (For more information about statewide testing, see https://education.mn.gov/MDE/fam/tests/; for the current list of MN's academic standards in mathematics, see https://education.mn.gov/MDE/dse/stds/Math/). The MCA-II mathematics is administered once a year in grades 3 to 8 and 11. Statewide MCA-II mathematics data were requested from MDE.

Study sample. The same cohort as the HSLS was used for the analysis, 9th graders in the 2009-10 school year. The sample size includes a total of 3,501 Asian

[2] MDE replaced using the MCA-II for grades 3 to 8 in 2010-11 and grade 11 in 2013-14 with the MCA-III.

students, 5,041 Black/African American students, and 48,384 White students. Descriptive data are provided in Tables 5 and 6.

Table 5

Frequencies of student demographics from the MCA data set.

Variable	Frequency (n)	Percent (%)
Race		
Asian	3,501	6.2
Black	5,041	8.9
White	48,384	85.0
Asian language subgroups		
English	1,007	28.8
Chinese-Mandarin	129	3.7
Hmong	1,509	43.1
Khmer	112	3.2
Lao	119	3.4
Vietnamese	250	7.1
Other Southeast Asian language	166	4.7
Other East Asian language	57	1.6
South Asian language	92	2.6
Other language	60	1.7
Sex		
Male	29,067	51.1
Female	27,859	48.9
English learner status		
English learner	2,022	3.6
Non-English learner	54,904	96.4
Free or reduced-priced lunch eligibility		
Qualifies for free or reduced-priced lunch	15,965	28.0
Does not qualify for free or reduced-priced lunch	40,961	72.0
Special education services		
Receives special education services	6,326	11.0
Does not receive special education services	50,600	88.9

Table 6
Descriptive statistics of MCA standardized math scores by grade.

Statistic	8th Grade	11th Grade
N	52,557	52,404
Minimum	14.20	24.53
Maximum	81.69	78.59
Mean	50.0	50.0
Median	50.7	50.5
Standard Deviation	10.0	10.0
Variation	100.0	100.0
Range	67.5	54.1
Kurtosis		
Statistic	0.9	0.2
Standard Error	0.0	0.0
Skewness		
Statistic	-0.2	-0.2
Standard Error	0.0	0.0

Variables. Demographic variables are typically reported by parents or guardians when students enroll in a MN school district. Race is a categorical, nominal variable where 1 = Asian, 2 = Black, and 3 = White. Primary home language is used as a proxy to identify Asian subgroups since detailed race and ethnic subgroup data are not collected by MDE[3]. In the 2011-12 school year, there were over 70 different languages spoken primarily at home among MN students. Language groups were recoded to identify Asian subgroups. Of 35 different languages spoken at home among Asian students, an Asian language subgroup variable was recoded, where 1 = English, 2 = Chinese-Mandarin, 3 = Hmong, 4 = Khmer, 5 = Lao, 6 = Vietnamese, 7 = Other Southeast Asian language (including Burmese[4], Cebuano, Indonesian, & Thai), 8 = Other East Asian language

[3] As part of MN's data disaggregation bill, detailed racial and ethnic group data for students of color will be collected in the 2018-19 school year in three school districts and two charter schools (see https://www.leg.state.mn.us/docs/2018/mandated/180150.pdf). All school districts and charter schools will be required to collect these data in the 2019-20 school year.

[4] It is very likely that this group contains mostly students with a Karen home language from Myanmar, formally known as Burma. At the time, however, Karen was not yet an available home language option. Of the 261 home language groups in MN, Karen was the 6th largest home language group overall and the 3rd

(including Japanese, Mongolian, Korean, and Tibetan), 9 = South Asian language (including Bengali, Hindi, Pashto, and Nepali), and 10 = Other language (e.g., Arabic, No specified language, English-Creolized, Spanish, and etc.). Sex is a dichotomous, nominal variable, where 1 = male and 2 = female. English Language Learner status is a proxy for English language proficiency, where 0 = non-English learner and 1 = English language learner. The ACCESS (Assessing Comprehension and Communication in English State-to-State) for ELLs test is used to determine students' proficiency in English and need for support services. Free or reduced-price lunch status (FRPL) is used as a proxy for family socioeconomic status. Students are eligible for FRPL through the National School Lunch program if their family's household income meets 185% of U.S. Census poverty threshold (see https://www.federalregister.gov/documents/2018/05/08/2018-09679/child-nutrition-programs-income-eligibility-guidelines). FRPL status is a nominal and categorical variable with 1 = does not qualify for FRPL and 2 = qualifies for FRPL. Equivalent variables for immigrant generation and first-generation college student status were not available. An additional variable used in the analysis of MCA-II data that was not available with the HSLS data is status of receiving special education services (SPED). It is important to control for this variable because students receiving SPED services have different educational needs, supports, and experiences. SPED services status is a nominal and categorical variable with 0 = does not receive SPED services and 1 = receives SPED services. Math coursework in spring 2012 from the MN Common Course Catalogue were available for students, but not used for the analysis for a couple reasons: 1) coursework

largest Asian language group during the 2017-18 school year (https://w20.education.state.mn.us/MDEAnalytics/Data.jsp). Burmese ranked the 38th largest home language group overall.

data is only available starting in 2010-11 school year meaning 8^{th} grade math coursework is not available given the cohort of interest, and 2) coursework data was only available for 30% of the students resulting in an unrepresentative sample. Math achievement is measured by the MCA-II scale score in 8^{th} grade and 11^{th} grade, as there are not any statewide assessments completed in 9^{th} grade. The MCA-II scale scores are a three-digit number with the first digit being the grade level, then a two-digit score. Only the two-digit score was retained for analysis. A score of 50 is the cut off for students to meet the proficiency standards, but proficiency levels were not examined for the purposes of this study. Scores range from 0 to 99, but are on different scales for each grade. Thus, the scores were standardized to t-scores using the mean score at each grade. The standardized scores have a mean of 50 and standard deviation of 10.

Statistical analysis and models. The following outlines the method, models, and statistical tests used to examine each research question. Statistical hypothesis testing will use an alpha level of $\alpha = 0.050$.

1) To what extent are the distribution of math scores for Asian, White, and Black student groups bimodal/multimodal? If there is evidence of multimodality, how to do the mode groups differ across demographics?

Method:
1) Run Silverman's Test of Multimodality and examine density plots of the distribution of math scores.
2) Examine mode groups by completing crosstabs with chi-square tests of independence and z-tests

Statistical tests: Silverman's Test for Multimodality

H_o: $k \leq x$
H_1: $k > x$
where k = number of modes and x = 1, 2, or 3

Chi-square Test of Independence or Fisher's Exact Test for 2x2 crosstabs

H_o: Mode group is independent of demographic variable

H_1: Mode group is not independent of demographic variable

Z-test for column proportions adjusting for p-values using the Bonferroni method

2) To what extent are there differences in mean math scores across AA subgroups?

Method: Run a one-way analysis of variance (ANOVA) with post-hoc tests comparing mean math scores across AA language groups.

Model specification: $Y_{ij} = \mu_i + \varepsilon_i$

where:

Y_{ij} = predicted math score for the *j*th student in the *i*th Asian language group

μ_i = mean math score for the *i*th Asian language group

ε_i = error term for the *i*th Asian language group

Statistical tests: ANOVA F-test for:

H_o: μ_i are equal

H_1: two or more μ_i are not equal

Tukey or Games-Howell post-hoc tests depending on whether data meet homogeneity of variance assumption

3) To what extent are there differences in math scores during high school by race while grouping Asians an aggregated group compared to disaggregating Asians into multiple subgroups?

Method: Use LME modeling to run a random intercept model to examine changes in math achievement between 8^{th} and 11^{th} with the following variables: sex, English learner status, eligibility for FRPL, and receipt of special education services.

- Race with an aggregated Asian group or disaggregated Asian subgroups (separate models were run with White students and Black students as the reference group for comparison with Asian subgroups)
- Sex (reference group = male)
- English learner status (reference group = non-English learner)
- Eligibility for FRPL (reference group = does not qualify for FRPL)
- Special education status (reference group = does not receive special education services)

Model specification:

Level 1: Within-subjects

$$Y_{ij} = \beta_{0j} + \beta_1(Grade_{ij}) + \beta_n(X_{nij}) + e_{ij}$$

where:

i =	individual
j =	time
Y_{ij} =	predicted math score for the *i*th individual at time *j*
β_{0j} =	mean math score for individuals at the starting grade level, 8th grade
β_1 =	rate of change for the *i*th individual, i.e., slope
β_n =	beta coefficient for *n*th covariate
X_{nij}=	*n*th covariate for the *i*th individual at time *j*
e_{ij} =	error term for the *i*th individual at time *j*

Level 2: Between-subjects

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

where:

γ_{00} =	mean math score for the whole sample, i.e., grand mean
μ_{0j} =	error for random intercepts for time *j*

Full Model:

$$Y_{ij} = (\gamma_{00} + \mu_{0j}) + \beta_{1j}(Grade_{ij}) + \beta_{2j}(X_{2ij}) + \cdots + \beta_{nj}(X_{nij}) + e_{ij}$$

Chapter 4: Results

Comparison of the Distribution of Math Scores across Racial Groups

HSLS. Asian students had a higher mean math score than White or Black students in both 9th and 11th grade (Table 7). Math scores for Asian students ranged between 24.9 to 82.2 with a mean score of 57.3 in 9th grade, and scores ranged from 28.9 to 82.9 with a mean score of 57.7 in 11th grade. Math scores in 9th grade ranged between 24.1 to 82.2 with a mean of 52.1 for White students, while 11th grade math scores ranged from 25.3 to 84.9 with a mean of 52.0. Black students had math scores ranging from 24.4 to 76.8 in 9th grade with a mean of 45.6, while in 11th grade scores ranged from 26.5 to 79.9 with a mean of 45.9. Asian students had higher standard deviations of math scores across racial groups at both assessment points, suggesting that the scores for Asian students are more spread out and have greater variation than math scores for White or Black students (Table 7). The standard deviation of the math score was 10.4 in 9th grade and 10.5 in 11th grade for Asian students, while White students had a standard deviation of 9.5 and 9.7 respectively, and Black students had a standard deviation of 9.1 and 8.4 respectively.

Figure 1 shows histograms of the distribution of scores by race for both grade levels. The distribution of 9th and 11th grade math scores for Asian students was slightly negatively skewed ($G1$ = -0.3 and $G1$ = -0.2 respectively) with more students having higher scores than lower scores. Ninth grade scores among Asian students had a sharper peak and heavier tails than a normal distribution ($G2$ = 0.1), while 11th grade scores had a slightly flatter than normal peak with lighter tails ($G2$ = -0.5). Ninth grade scores appeared to be fairly unimodal, while 11th grade scores appeared to be bimodal with the sharpest and largest mode at a score of 52 and a slightly smaller mode at about 63.

Scores for Black students in 9th and 11th grade was slightly positively skewed (*G1* = 0.1 for both grades), where more students had lower scores than higher scores (Table 7 and Figure 1). The distributions of scores among Black students in both grade levels had flatter peaks and lighter tails than the normal distribution (*G2* = -0.4 for both grades). The histogram of 9th grade scores among Black students appeared unimodal, whereas 11th grade scores showed evidence of two similarly sized modes centered at about a score of 39 and 49. The distributions of scores for White students was slightly negatively skewed in 9th grade (*G1* = -0.1) and approximately normal in 11th grade (*G1* = 0.0) (Table 7 and Figure 1). Among White students, the distribution of 9th grade scores was normally peaked (*G2* = 0.0) while the distribution for the 11th grade scores had a slightly flatter peak and lighter tails (*G2* = -0.3).

Table 7
Descriptive statistics of HSLS math theta scores across race by grade.

Statistic	9th Grade			11th Grade		
	Asian	Black	White	Asian	Black	White
N	170	600	2,310	170	610	2,330
Minimum	24.9	24.4	24.1	28.9	26.5	25.3
Maximum	82.2	76.8	82.2	82.9	79.9	84.9
Mean	57.3	45.6	52.1	57.7	45.9	52.0
Median	58.0	45.8	51.8	56.4	46.9	51.8
Standard Deviation	10.4	9.1	9.5	10.5	8.4	9.7
Variance	108.2	82.6	90.6	110.4	71.0	94.2
Range	57.3	52.4	58.1	54.0	53.4	59.6
Skewness (G1)						
Statistic	-0.3	0.1	-0.1	-0.2	0.1	0.0
Standard Error	0.2	0.1	0.1	0.2	0.1	0.1
Kurtosis (G2)						
Statistic	0.1	-0.4	0.0	-0.5	-0.4	-0.3
Standard Error	0.4	0.2	0.1	0.4	0.1	0.1

Note. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

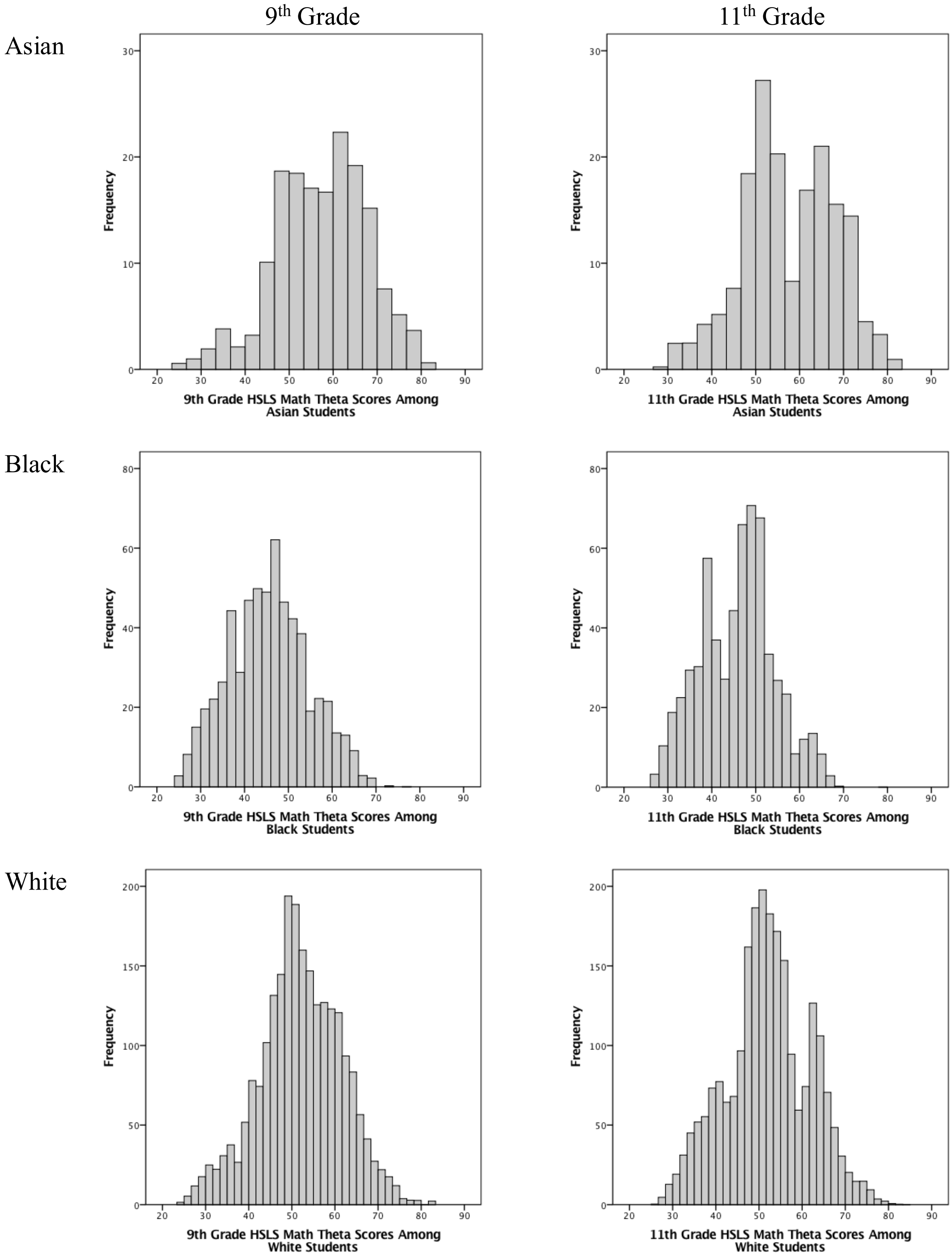
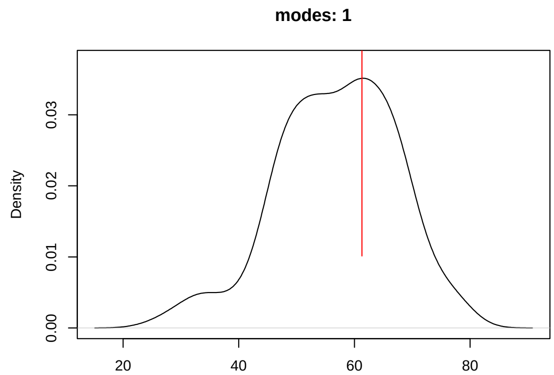


Figure 1. Histogram plots of the distribution of HSLS math theta scores by race and grade.

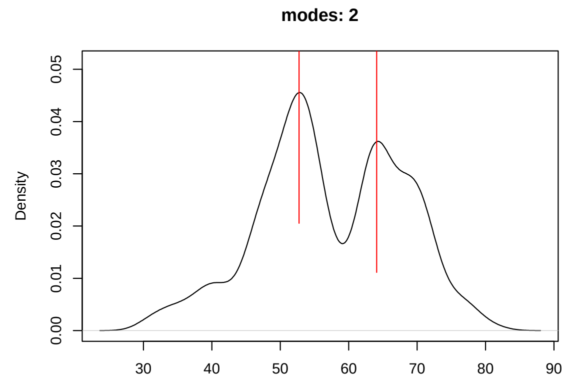
The Silverman Test was used to estimate the number of modes present in the distribution of scores. For White students, the distribution of 9th grade math theta scores appeared unimodal, while potentially three modes existed across 11th grade scores with the sharpest and largest mode at about at about a score of 50, a medium-sized mode at a score of roughly 63, and a smaller mode at a score of 40. of math scores across race and grade (Table 8 and Figure 2). Among Asian students, 9th grade scores were unimodal at a score of about 62, and 11th grade scores were bimodal with the largest mode at a score of 52 and somewhat smaller mode at 65. Results for Black students suggest the distribution of scores at both grade levels were multimodal with at least three modes. In 9th grade, two similarly sized large modes were identified at scores of about 44 and 47, while a third mode and slightly smaller mode was identified at a score of about 38. In 11th grade, the largest mode was identified at a score of about 49, a medium-sized mode at a score of about 39, and a small mode at a score of about 63. Among White students, 9th grade scores were unimodal at a score of 50 and 11th grade scores had at the most three modes, a large mode at a score of about 50, a medium-sized mode at a score of 63, and a slightly smaller mode at a score of about 40.

Asian

9th grade
N = 1,200 Bandwidth = 3.273

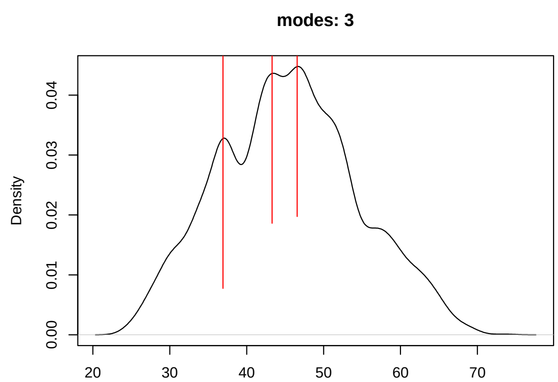


11th grade
N = 1,250   Bandwidth = 1.968

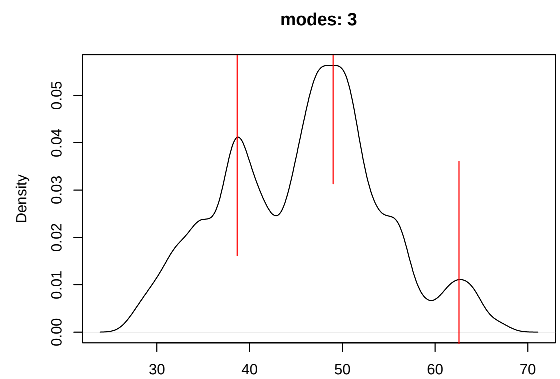


Black

9th grade
N = 5,400 Bandwidth = 1.359

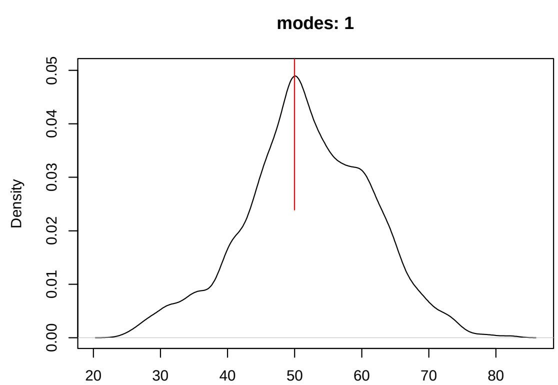


11th grade
N = 5,480 Bandwidth = 1.132



White

9th grade
N = 19,340 Bandwidth = 1.259



11th grade
N = 19,500 Bandwidth = 0.962

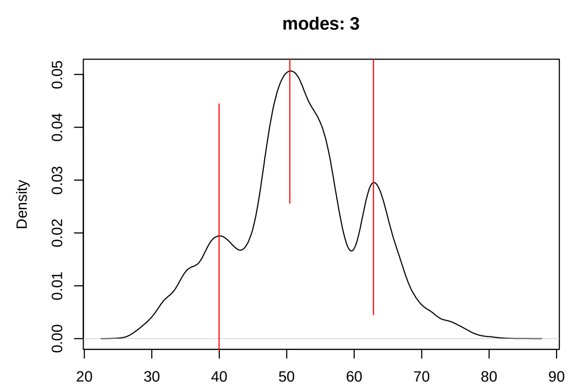


Figure 2. Density plots of the distribution of HSLS math theta scores by race and grade with identified modes (up to three) and minimum bandwidth from Silverman's Test of Multimodality. Modes are identified by the vertical lines.

Table 8

Silverman Test of Multimodality results of HSLS math theta scores by race and grade.

Race by Grade	Number of modes tested	*p*-value	Minimum bandwidth
Asian			
9th Grade (N = 1,200)	1	0.071†	3.273
11th Grade (N = 1,250)	1	0.000***	
	2	0.106	1.968
Black			
9th Grade (N = 5,400)	1	0.006**	
	2	0.003**	
	3	0.007**	
11th Grade (N = 5,480)	1	0.000***	
	2	0.003**	
	3	0.000***	
White			
9th Grade (N = 19,340)	1	0.124	1.259
11th Grade (N = 19,500)	1	0.001**	
	2	0.001**	
	3	0.292	0.962

Note. Since the HSLS data were weighted to account for the complex sampling design, new data samples were created with cases replicated by their weight multiplied by a factor of 10. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

†$p < .100$, *$p < .050$, **$p < .010$, ***$p < .001$

Where applicable, differences in demographics across mode groups were examined with crosstabs, Chi-square Tests of Independence or Fisher's Exact Tests, and z-tests of column proportions (Figures 3 to 6). Among Asian students, Silverman's Test results showed that 9th grade math theta scores were unimodal so mode groups were not examined. At 11th grade, math theta scores were bimodal. Asian students in Mode Group 1 had scores less than 60 (N = 100), while those in Mode Group 2 had scores of 60 or greater (N = 80). Using an alpha level of $\alpha = 0.010$, mode groups significantly differed by 185% poverty threshold, most advanced math course taken, and advanced math course-taking (Figure 3). More students in Mode Group 1 were from households with incomes below the 185% poverty threshold than those in Mode Group 2, 45% compared to 23% respectively. Most students in Mode Group 1 took Algebra II/III (45%) or pre-calculus and/or analysis and functions (29%). Most students in Mode Group 2 took pre-calculus and/or analysis and functions (42%) as well, but they were more likely to take calculus than students in Mode 1—37% compared to 2% respectively. Students in Mode Group 2 (51%) were also more likely to have taken an advanced math course than those in Mode Group 1 (15%). The mode groups had statistically similar proportions of students across categories of sex, immigrant generation, first language learned, and U.S. region. Differences across Asian subgroup and first-generation college student status were borderline statistically significant with p-values close to $\alpha = 0.050$. Mode Group 1 had more Southeast Asian students (37% compared to 21% of students in Mode 2), whereas Mode Group 2 had more Chinese students (32% compared to 16% of students in Mode 1). While most students in both mode groups would not be the first in their family to attend college (68% of students in Mode Group 1 and 70% of students in Mode Group 2), there were more students in Mode Group 1 who would be first-generation college students (31%) than those in Mode Group 2 (16%). Between the two modes, there were similar proportions of students across sex (51% compared to 49% respectively within Mode Group 1 and 48% compared to 52% respectively within Mode Group 2). Over two-thirds of students in both modes were second-generation immigrants (68% of those in Mode Group 1 and 70% of those in Mode 2), while one-quarter of the students were first-generation immigrants (25% of those in Mode Group 1 and 24% of those in Mode Group 2). Nearly half of the students in both mode groups were from Western U.S. compared to other regions, 48% of students in Mode Group 1 and 44% of students in Mode Group 2.

At both 9th and 11th grade, the distribution of math theta scores was multimodal among Black students. In 9th grade, Mode Group 1 consisted of students with scores less than 39 (N = 150), Mode Group 2 included students with scores between 39 to 46 (N = 160), and Mode Group 3 included students scoring higher than 46 (N = 300). In 11th grade, Mode Group 1 included students with scores less than 44 (N = 240), Mode Group 2 with students scoring between 44 and 60 (N = 340), and Mode Group 3 with students scoring higher than 60 (N = 40). In 9th grade, mode groups differed significantly at the $\alpha = 0.010$ level by all characteristics with the exception of immigrant generation (Figures 4 and 5). In 11th grade, mode groups differed significantly by first-generation college student status, 185% poverty threshold, and most advanced math course taken. More students in the low scoring Mode Group 1 in 9th grade were male students, 54% compared to 35% of students in Mode Group 2 and 41% of students in Mode Group 3. While nearly all students in all three modes had an English first language, there were more students in Mode Group 3 with an English first language at 9th grade, 96% compared to 90% of students in Mode Group 1 and 89% of students in Mode Group 2. At both grade levels, Mode Group 1 consisted of more students who would be first-generation college students than mode groups 2 and 3 (49% compared to 26% for both mode groups 1 and 2 at 9th grade, 42% compared to 28% of Mode Group 2 and 15% of Mode Group 3 at 11th grade). There were more students from Southern U.S. in the higher scoring Mode Group 3 at 9th grade, 64% compared to 50% of those in Mode Group 1 and 57% of students in Mode Group 2. The majority of the students in Mode Group 1 were from households with incomes below the 185% poverty threshold than those in mode groups 2 and 3, 68% compared to 60% of Mode Group 2 and 50% of Mode Group 3 in 9th grade and 64% compared to 55% of Mode Group 2 and 35% of Mode Group 3 in 11th grade). At 9th grade, more students across mode groups took Algebra II/III than other types of courses, however, a higher proportion of students in Mode Group 2 did so—58% compared to 38% of Mode Group 1 and 46% of Mode Group 3. Students in Mode Group 1 (30%) were more likely to take geometry than those in mode groups 2 (11%) and 3 (7%). Additionally, Mode Group 3 students were more likely to take pre-calculus or analysis and functions than students in the other two groups (20% compared to 3% of Mode Group 1 students and 2% of Mode Group 2 students). At 11th grade, students in mode groups 1 and 2 were more likely to take algebra II/III (44% and 50% respectively compared to 38% of Mode Group 3), whereas students in Mode Group 3 were more likely to take pre-calculus or analysis and functions (44% compared to 3% of Mode Group 1 and 12% of Mode Group 2). Over three-quarters of the students in all three modes at 9th grade had not ever taken advanced math courses (between 78% to 93%), but students in Mode Group 3 (21%) were more likely to do so then mode groups 1 (7%) and 2 (9%). Mode groups did not differ significantly at either grade level on immigrant generation, where at least two-thirds of students were third-generation or later (between 70% to 85%).

Among White students, only the distribution of 11th grade scores showed the presence of multiple modes. Mode Group 1 included students scoring less than 43 (N = 420), Mode Group 2 had students scoring between 43 and 60 (N = 1,380), and Mode Group 3 consisted of students with scores higher than 60 (N = 520). Mode groups significantly differed across first-generation college student status, U.S. region, 185% poverty threshold, most advanced math course taken, and advanced math course-taking (Figure 6). While most students across mode groups would not be the first person in their families to go to college (between 64% to 88%), there more students in the lower scoring

Mode Group 1 who would be first-generation college students compared to the other modes, 37% compared to 26% of Mode Group 2 and 12% of Mode Group 3. Mode Group 1 students were more likely to live in Southern U.S., 41% compared to 34% of Mode Group 2 and 28% of Mode Group 3. Mode Group 1 (45%) also included more students from households with incomes below the 185% poverty threshold than mode groups 2 (27%) and 3 (17%). Students in mode groups 1 and 2 were most likely to take algebra II/III (44% and 48% respectively compared to 19% of Mode Group 3), while most students in Mode Group 3 took pre-calculus or analysis and functions (49% compared to 3% of Mode Group 1 and 16% of Mode Group 2. Mode groups had similar proportions across immigrant generations and first language learned, where the vast majority of students in all three groups were third-generation or later (89% to 94%) and had an English first language (97% to 99%). There were slightly more male students in mode groups 1 (54%) and 3 (55%) compared to Mode Group 2 (50%), but the difference was only borderline statistically significant.

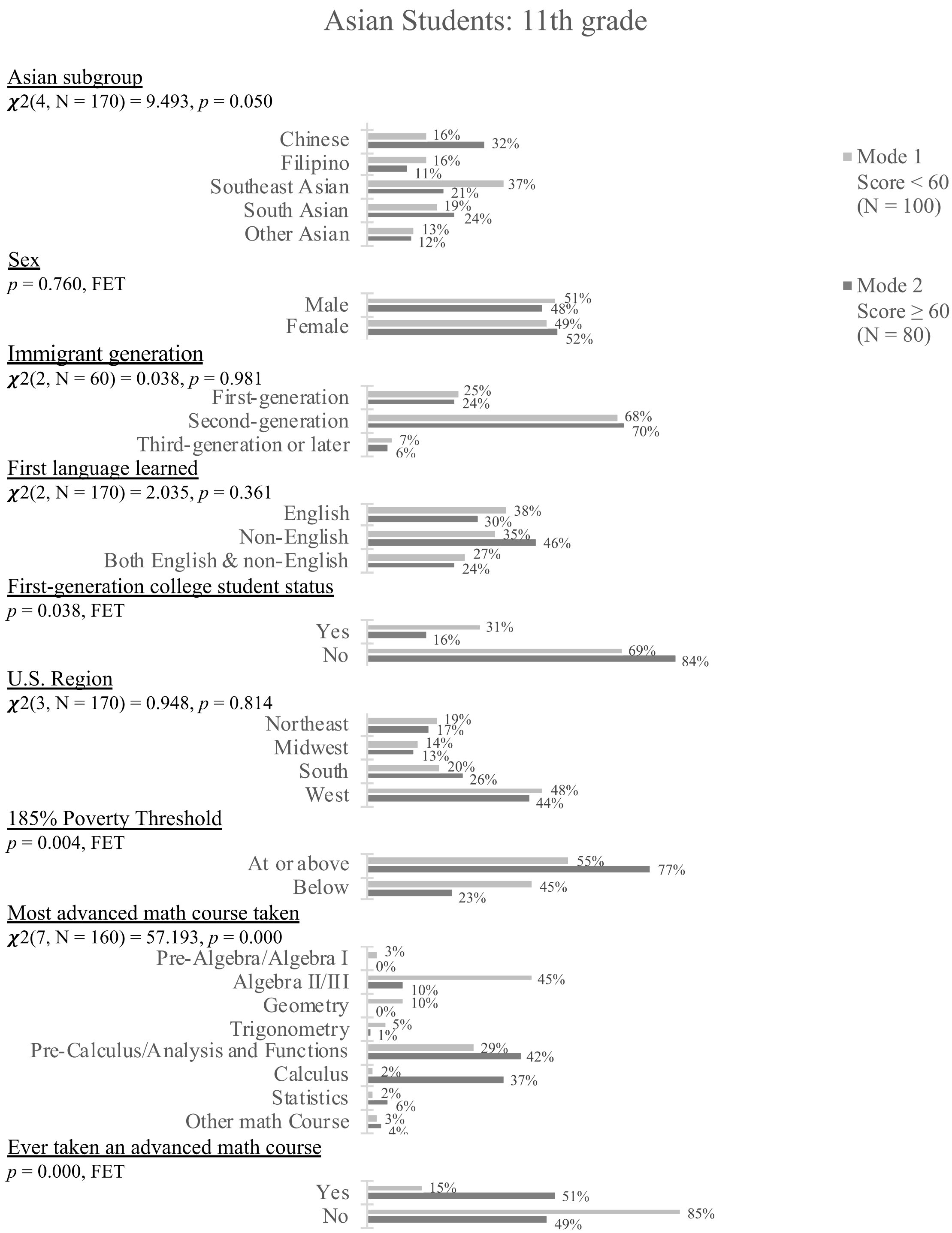


Figure 3. Mode groups of HSLS math theta scores by demographics in 11th grade for Asian students. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

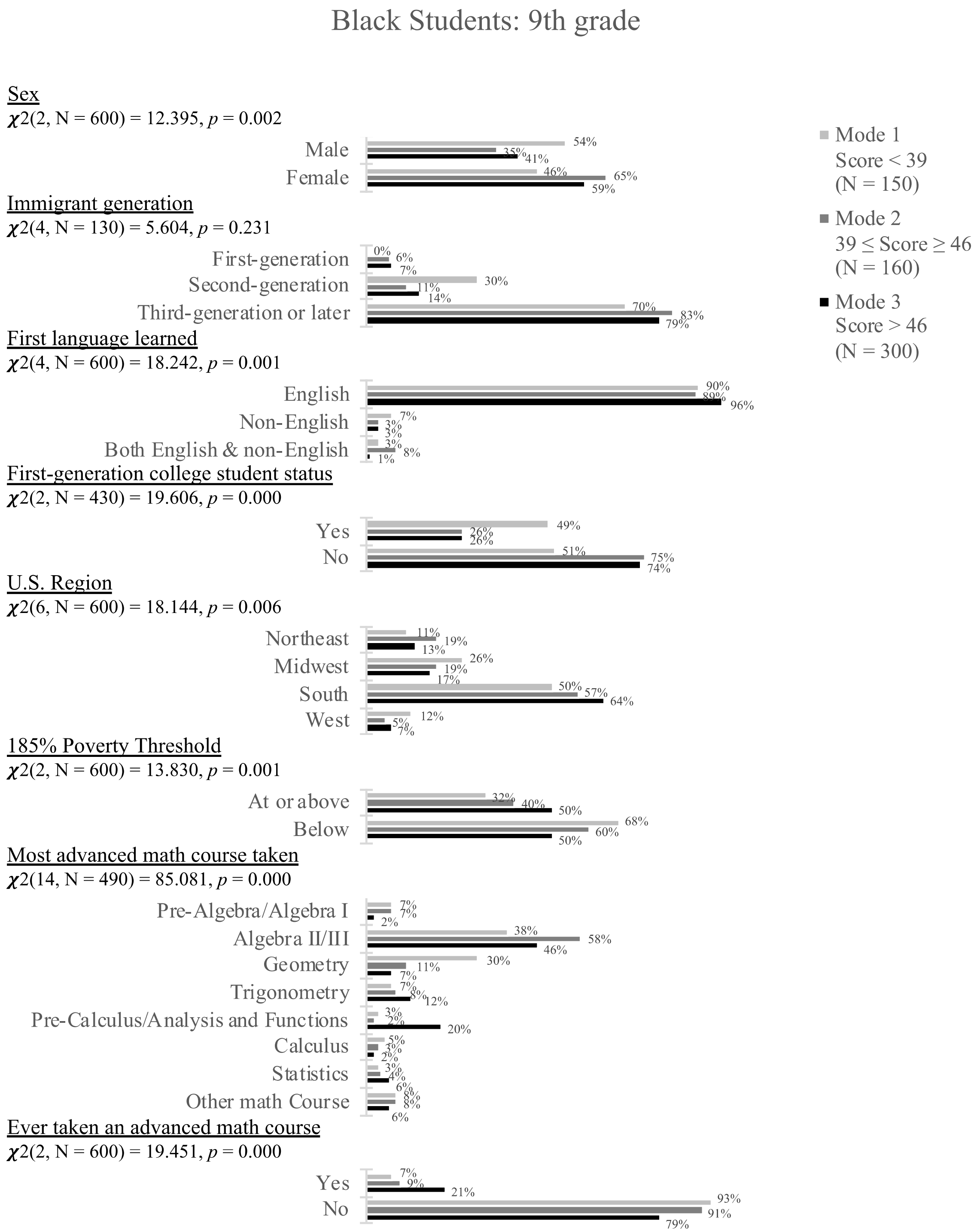


Figure 4. Mode groups of HSLS math theta scores by demographics in 9th grade for Black students. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

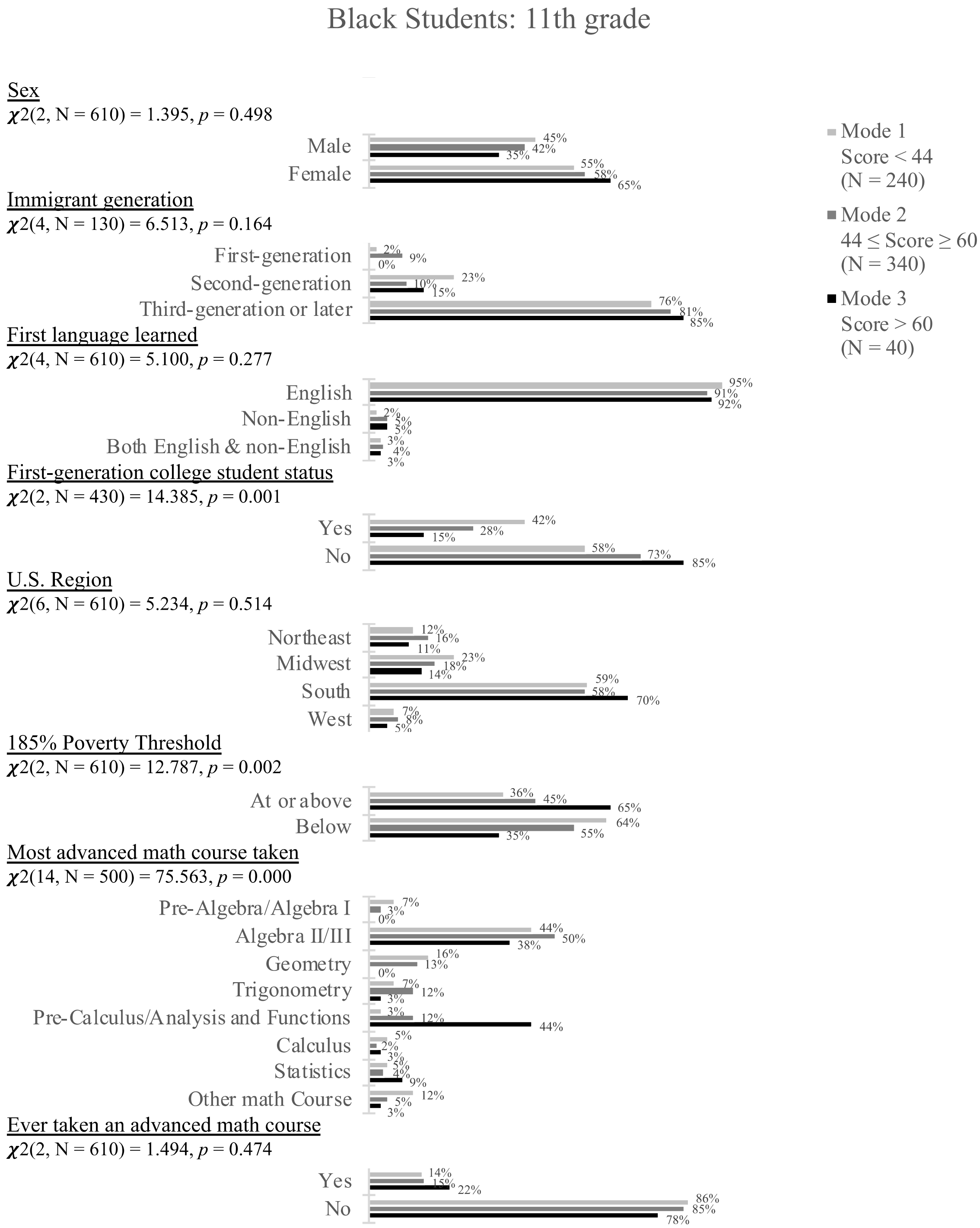


Figure 5. Mode groups of HSLS math theta scores by demographics in 11th grade for Black students. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

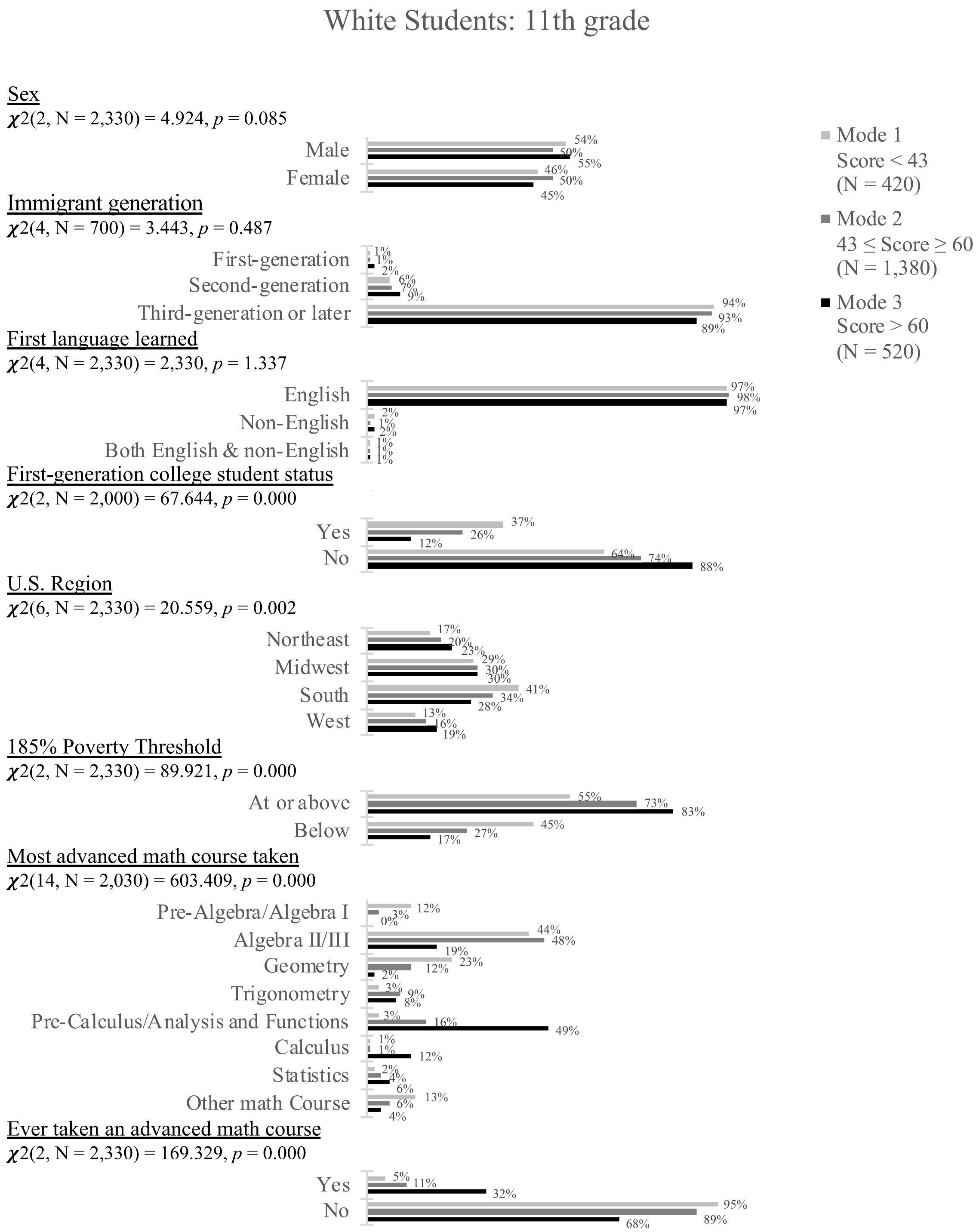


Figure 6. Mode groups of HSLS math theta scores by demographics in 11th grade for White students. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

MCA. Asian students had slightly lower mean math scores than White students in 8th grade and 11th grade, but had higher mean math scores than Black students at both assessment points (Table 9). Mean math scores for Asian students were 50.3 in 8th grade and 49.1 in 11th grade, whereas White students had mean math scores of 50.8 and 51.0 respectively by grade and Black students had mean math scores of 41.0 and 40.3 respectively by grade. The range of scores in 8th grade was the same for Asian students and White students (14.2 to 81.7), while Black students had a slightly narrower range of scores between 14.2 to 78.3. In 11th grade, all three groups had the same range of scores between 24.5 to 78.6. Further, Asian students had slightly higher standard deviations of math scores across racial groups at both assessment points (Table 9). In the MCA data, the standard deviation for math scores was 10.5 in 8th grade and 10.6 in 11th grade for Asian students, while White students had a standard deviation of 9.5 for both grades and Black students had a standard deviation of 10.1 and 9.6 respectively by grade level.

Table 9
Descriptive statistics of MCA standardized math scores across race by grade.

Statistic	8th Grade			11th Grade		
	Asian	Black	White	Asian	Black	White
N	2,756	4,266	45,535	3,290	4,159	44,955
Minimum	14.2	14.2	14.2	24.5	24.5	24.5
Maximum	81.7	78.3	81.7	78.6	78.6	78.6
Mean	50.3	41.0	50.8	49.1	40.3	51.0
Median	50.0	41.1	51.4	48.8	40.5	51.0
Standard Deviation	10.5	10.1	9.5	10.6	9.6	9.5
Variance	109.7	102.9	90.9	112.6	92.9	90.0
Range	67.5	64.1	67.5	54.1	54.1	54.1
Skewness (G1)						
Statistic	0.0	-0.2	-0.2	0.0	0.2	-0.1
Standard Error	0.0	0.0	0.0	0.0	0.0	0.0
Kurtosis (G2)						
Statistic	0.7	0.2	1.0	0.1	-0.4	0.3
Standard Error	0.1	0.1	0.0	0.1	0.1	0.0

Histograms were run to examine the distribution of MCA scores for each racial group at both grade levels before performing any statistical tests (Figure 7). The distribution of MCA 8th and 11th grade scores for Asian students were symmetric (*G1* = 0.0 for both grades). The distributions of scores has slightly sharper peaks and heavier tails than a normal distribution (*G2*=0.7 for 8th grade scores and *G2*=0.1 for 11th grade scores). Ninth grade scores appeared to be unimodal at a score of 50, but 11th grade scores showed a presence of bimodality—a large mode centered at 48 and a very small mode centered at 24. Among Black students, the distribution of 8th grade scores were slightly negatively skewed (*G1* = -0.2) with a relatively normal peak, but slightly heavier tails (*G2* = 0.2). The distribution of 11th grade scores were slightly positively skewed with a normal peak and lighter tails (*G1* = 0.2 and *G2* = -0.4). There is a potential presence of bimodality at both grade levels. In 8th grade, the largest mode was at a score of 38, while a small mode appeared at a score of 14. In 11th grade, the largest mode was centered at 43, while another large but smaller mode was centered at 25. The distribution of MCA math scores for White students were slightly right skewed with sharper peaks and heavier tails than a normal distribution (*G1* = -0.2 and *G2* = 1.0 for 8th grade and *G1* = -0.1 and *G2* = 0.3 for 11th grade). Ninth grade scores appeared unimodal at a score of 53, while there was potential for bimodality across 11th grade scores with a large mode at a score of 48 and a very small mode at a score of 24.

Silverman's Test results suggest there are three or more modes at each grade level for each racial group (Table 10 and Figure 8). Among Asian students, the largest mode at 8th grade was identified at a score of about 50 with two small modes at a score of about 18 and 79. At 11th grade, the largest mode was centered at a score of 49 with two small

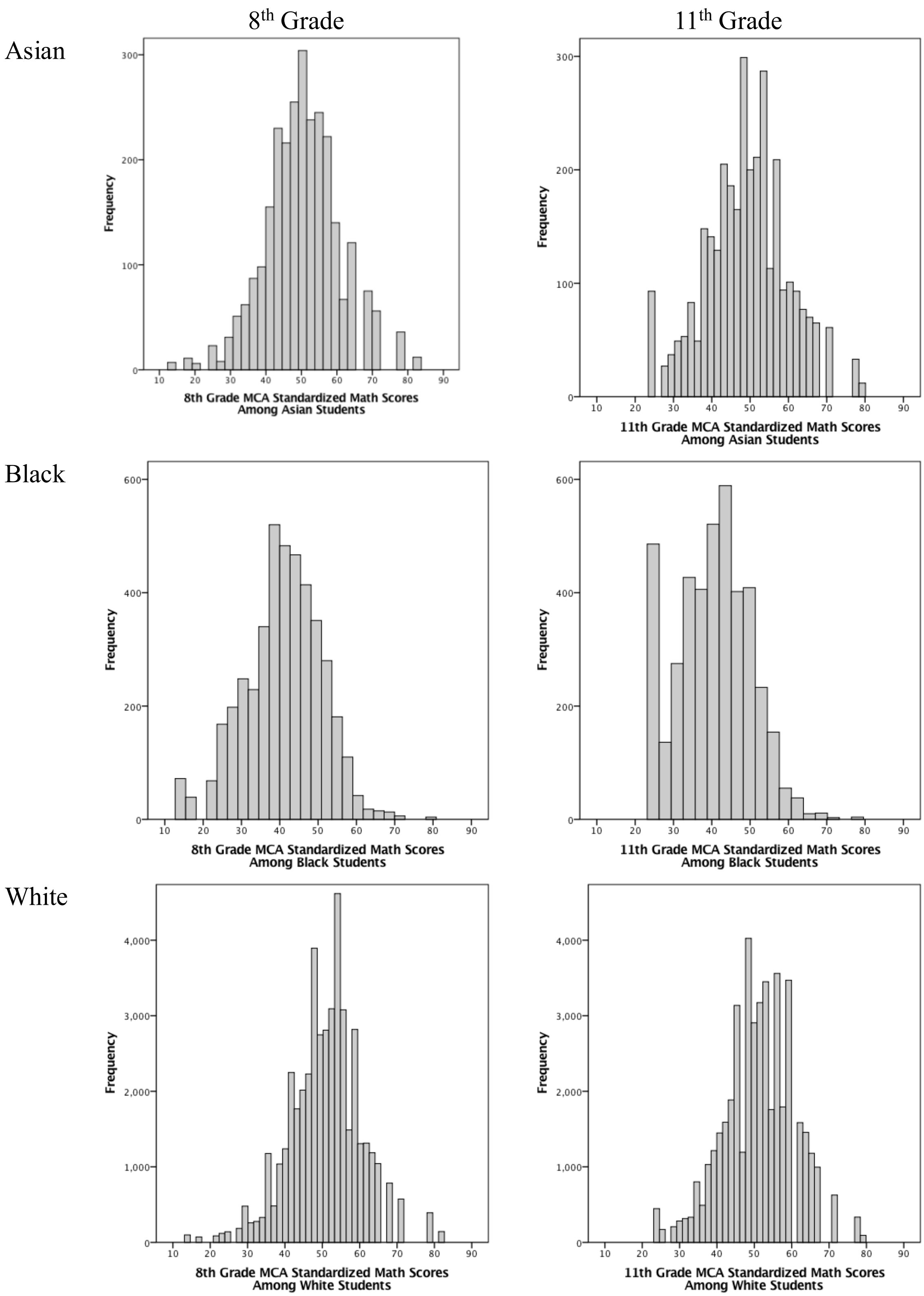


Figure 7. Histogram plots of the distribution of MCA standardized math scores by race and grade.

modes at a score of 25 and 78. For black students, the distribution of 8th grade scores had the largest mode identified at a score of 40, with two small modes at scores of about 14 and 79. At 11th grade, the largest mode was identified at a score of about 42, a medium-sized mode was identified at a score of about 25, and a very small mode was identified at a score of 79. As for White students, the largest mode across 8th grade scores was identified at a score of about 52 with two very small modes identified at scores of about 15 and 79. The distribution of 11th grade scores had the largest mode identified at a score of 50 with two very small modes at scores of about 25 and 78.

Table 10

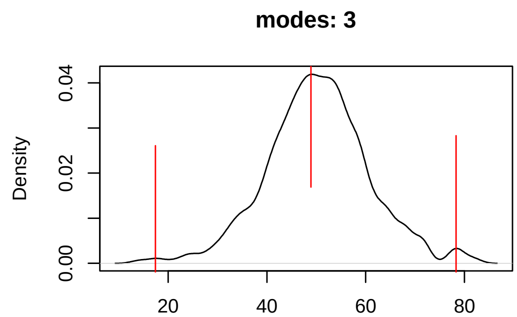
Silverman Test of Multimodality results of MCA standardized math scores by race and grade.

Race by Grade	Number of modes tested (k)	p-value	Minimum bandwidth
Asian			
8th Grade (N = 2,756)	1	0.000***	2.966
	2	0.040*	2.305
	3	0.191	1.642
11th Grade (N = 3,290)	1	0.000***	2.609
	2	0.000***	2.335
	3	0.000***	1.718
Black			
8th Grade (N = 4,266)	1	0.039*	2.725
	2	0.000***	2.518
	3	0.051†	1.708
11th Grade (N = 4,159)	1	0.022*	3.136
	2	0.000*	3.098
	3	0.516	1.307
White			
8th Grade (N = 45,535)	1	0.000***	3.040
	2	0.000***	2.345
	3	0.000***	1.546
11th Grade (N = 44,955)	1	0.000***	2.527
	2	0.000***	2.219
	3	0.000***	1.536

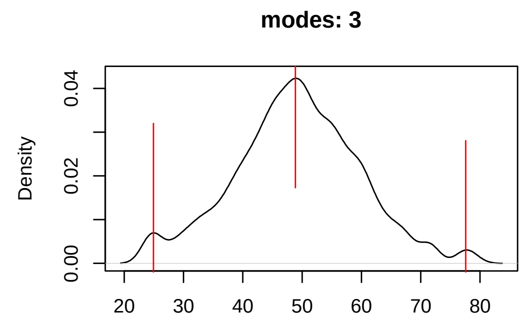
Note: When K = 1, an adjustment was used to improve the accuracy of the p-value (Hall & York, 2001).
†$p < .100$, *$p < .050$, **$p < .010$, ***$p < .001$

Asian

8th grade
N = 2,756 Bandwidth = 1.642

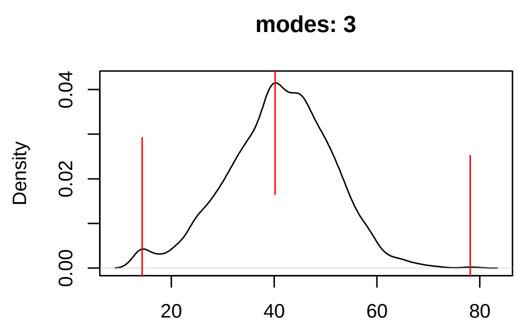


11th grade
N = 3,290 Bandwidth = 1.718

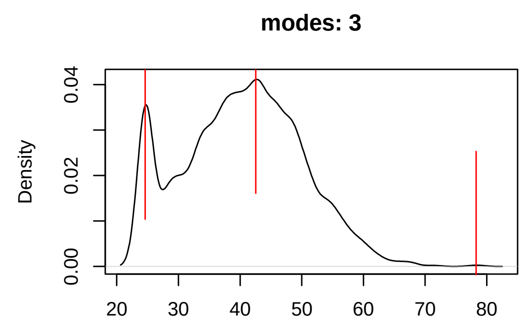


Black

8th grade
N = 4,266 Bandwidth = 1.708

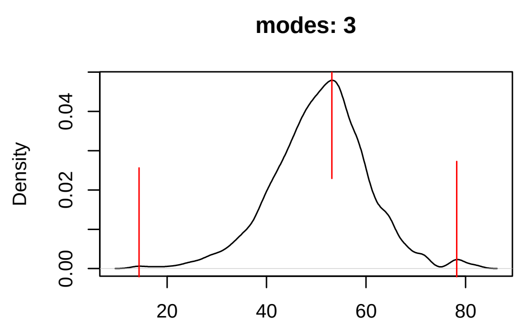


11th grade
N = 4,159 Bandwidth = 1.307



White

8th grade
N = 45,535 Bandwidth = 1.546



11th grade
N = 44,955 Bandwidth = 1.536

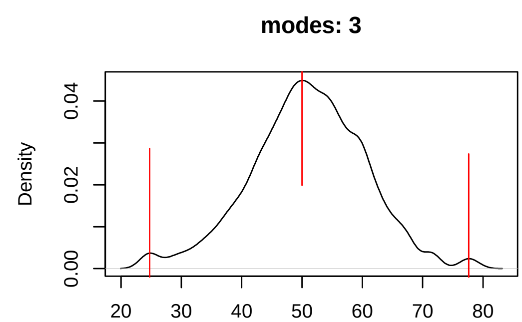


Figure 8. Density plots of the distribution of MCA standardized math scores by race and grade with identified modes (up to three) and minimum bandwidth from Silverman's Test of Multimodality.

Crosstabs with Chi-square Tests of Independence or Fisher's Exact Tests, and z-tests of column proportions were run to examine differences across mode groups for each racial group at both grade levels (Figures 9-11). Across racial groups, the number of students in the lowest (Mode Group 1) and highest (Mode Group 3) scoring mode groups were very small compared to the number of students in the mid-range (Mode Group 2) scoring mode group.

Among Asian students, Mode Group 1 had students with scores less than 18 (N = 18), Mode Group 2 with scores between 18 and 72 (N = 2,690), and Mode Group 3 (N = 48) with scores higher than 72 at 8^{th} grade. At 11^{th} grade, Mode Group 1 had students with scores less than 25 (N = 64), Mode Group 2 with scores between 25 and 77 (N = 3,181), and Mode Group 3 with scores greater than 77 (N = 45). Mode groups significantly differed by language subgroup, English learner status, eligibility for FRPL, and SPED services across the mode groups at both grade levels (Figure 9). Compared to the higher scoring mode groups 2 and 3, Mode Group 1 consisted mostly of students with a Southeast Asian home language (89% compared to 57% of Mode Group 2 and 10% of Mode Group 3 in 8^{th} grade; 83% compared to 62% of Mode Group 2 and 7% of Mode Group 3 in 11^{th} grade). Mode Group 3 had significantly higher proportions of students with an English home language (60% compared to 34% of Mode Group 2 and 6% of Mode Group 3 in 8^{th} grade; 60% compared to 28% of Mode Group 2 and 11% of Mode Group 3 in 11^{th} grade) and East Asian home language (23% compared to 5% of Mode Group 2 and 6% of Mode Group 3 in 8^{th} grade; 22% compared to 5% of Mode Group 2 and 2% of Mode Group 3 in 11^{th} grade) at both grades compared to mode groups 1 and 2. Mode Group 1 had significantly higher proportions of English learners (89% compared to

25% of Mode Group 2 and 3% of Mode Group 3 in 8th grade; 84% compared to 35% of Mode Group 2 and 4% of Mode Group 3 in 11th grade), students qualifying for FRPL (94% compared to 56% of Mode Group 2 and 19% of Mode Group 3 in 8th grade; 91% compared to 60% of Mode Group 2 and 16% of Mode Group 3 in 11th grade), and students receiving SPED services (56% compared to 6% of Mode Group 2 and 0% of Mode Group 3 in 8th grade; 27% compared to 5% of Mode Group 2 and 2% of Mode Group 3 in 11th grade) compared to mode groups 2 and 3. Differences in the proportion of male or female students across modes were not statistically different. However, Mode Group 1 had more male students compared to mode groups 2 and 3 in 8th grade (61% compared to 48% for both mode groups 2 and 3), while Mode Group 3 had more male students than mode groups 1 and 2 in 11th grade (67% compared to 47% and 49% respectively).

Since primary home language was available for all students in the MCA dataset, it was used to examine differences across mode groups for Black students and White students. Black students were grouped into four subgroups, English (N = 3,891, 77%), Somali (N = 617, 12%), Other African language (N = 330, 6.5%), and Other language (N = 203, 4%). There were two subgroups for White students, English (N = 47, 855, 99%) and Other language (N = 529, 1%).

Among Black students, only two modes were compared because the highest performing mode had very small sample sizes of 10 or less students at both grade levels. Students in this mode group were combined with students in Mode Group 2. In 8th grade, Mode Group 1 (N = 72) included students score less than 17, while Mode Group 2 included students scoring 17 or higher (N = 4,194). In 11th grade, Mode Group 1 (N =

352) included students with scores less than 25, while Mode Group 2 (N = 3,807) included students with scores of 25 or higher. Mode groups had significantly different proportions across language subgroup, English learner status, eligibility for FRPL, and SPED services at both grade levels (Figure 10). The large majority of students in both mode groups had an English home language, 75% for Mode Group 1 and 82% for Mode Group 2 at 8th grade and 72% for Mode Group 1 and 75% for Mode Group 2. At both grades, Mode Group 1 had a higher proportion of students with a Somali home language than Mode Group 2 (17% compared to 8% respectively at 8th grade and 21% compared to 13% respectively at 11th grade). While most students across mode groups were not English learners (between 75% to 95%), Mode Group 1 had significantly more students who were English learners compared to those who were not (17% compared to 5% of Mode Group 2 in 8th grade; 24% compared to 14% of Mode Group 2 in 11th grade). The majority of students in both mode groups qualified for FRPL in 8th and 11th grade (between 74% to 89%), but more students in Mode Group 1 were eligible than those in Mode Group 2 (86% compared to 74% respectively in 8th grade; 89% compared to 74% respectively in 11th grade). More students in Mode Group 1 received SPED services than those in Mode Group 2 (63% compared to 19% respectively in 8th grade; 41% compared to 12% respectively in 11th grade). The mode groups had similar proportions of male students and female students at about half and half respectively at both grade levels.

For White students, Mode Group 1 consisted of scores less than 21 (N = 170), Mode Group 2 with scores between 21 and 72 (N = 44,830), and Mode Group 3 (N = 535) with scores higher than 72 at 8th grade. In 11th grade, Mode Group 1 consisted of scores less than 26 (N = 617), Mode Group 2 with scores between 26 and 77 (N =

44,830), and Mode Group 3 with scores higher than 77 (N = 427). Mode groups significantly differed by language subgroup, English learner status, eligibility for FRPL, and SPED services at both grade levels, as well as by sex in 11^{th} grade (Figure 11). White students had very large sample sizes so statistical tests are likely to be statistically significant. While z-tests results indicated significant differences in proportions in language subgroup and English learners across mode groups, all or almost all students had an English home language (98% for Mode Group 1 and 99% for both mode groups 2 and 3 at both grade levels) and were not English learners (99% for both mode groups 1 and 2 and 100% for Mode Group 3 at both grade levels). Proportions of male students and female students were similar across mode groups in 8^{th} grade, but significant differences existed in 11^{th} grade where Mode Group 3 (71%), the highest performing group, had more male students than mode groups 1 (58%) and 2 (50%). Also, in both grades, Mode Group 1 had more students eligible for FRPL services than mode groups 2 and 3 (43% compared to 21% and 5% respectively in 8^{th} grade; 51% compared to 19% and 1% respectively in 11^{th} grade). Students in Mode Group 1 were also more likely to receive SPED services, 91% compared to 11% of Mode Group 2 and 3% of Mode Group 3 in 8^{th} grade, and 61% compared to 8% of Mode Group 2 and 3% of Mode Group 3 in 11^{th} grade.

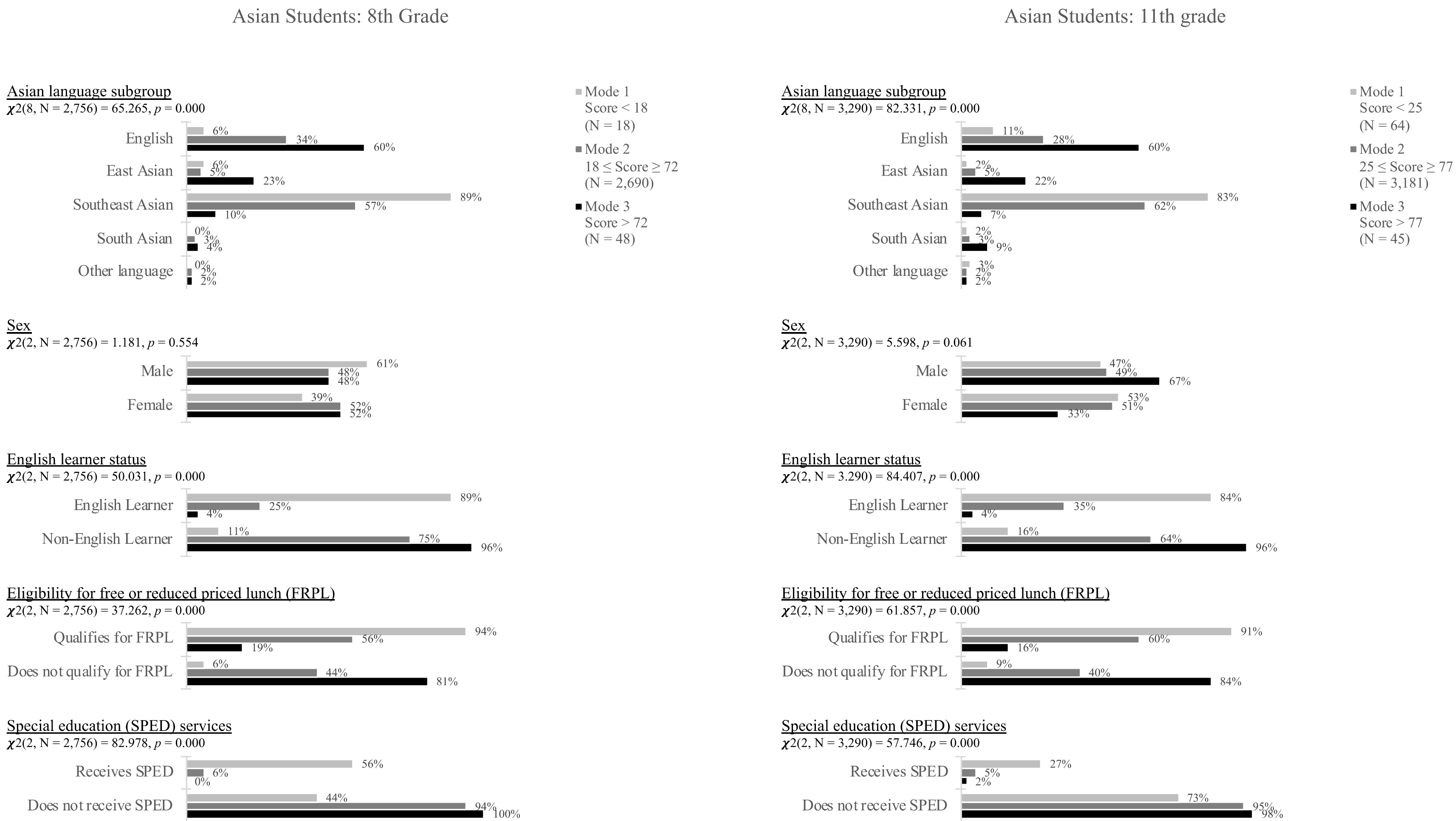
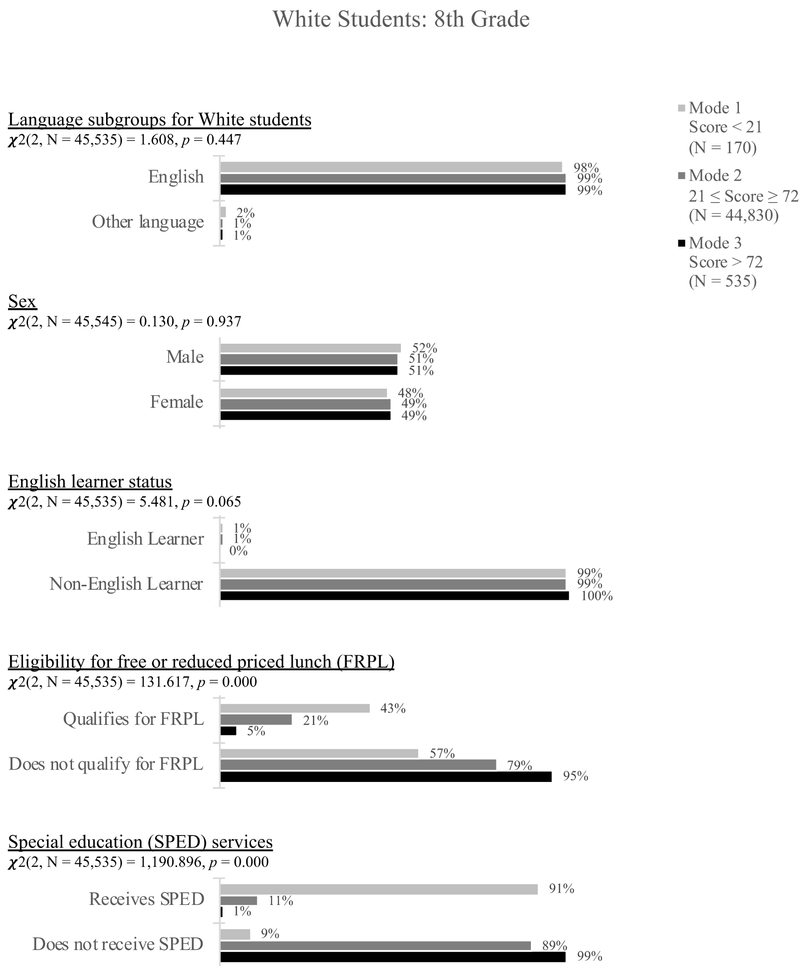


Figure 9. Mode groups of MCA standardized math scores by grade and demographics for Asian students. Asian subgroups were recategorized to reduce the percentage of cells with expected counts of less than five.

Figure 10. Mode groups of MCA standardized math scores by grade and demographics for Black students. Given the sample size of students in the highest scoring mode (Mode 3) was very small (N ≤ 10), Mode 3 was combined with Mode 2.



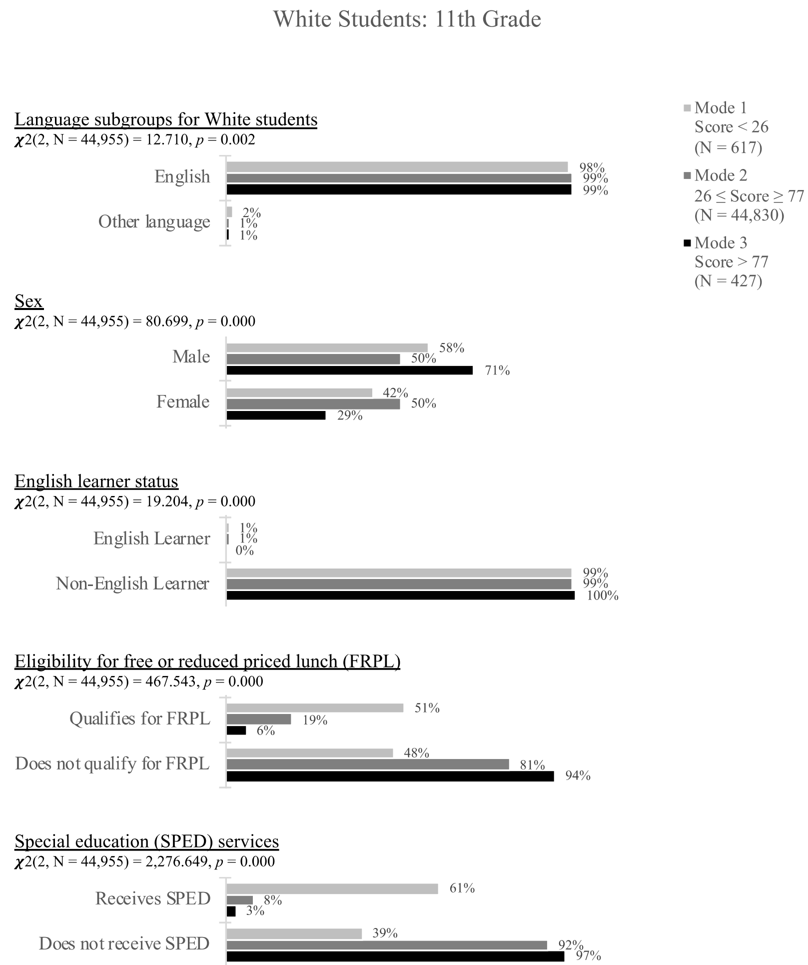


Figure 11. Mode groups of MCA standardized math scores by grade and demographics for White students.

Comparison of Math Achievement across Asian Subgroups

HSLS. Box plots of math scores across Asian subgroups are shown in Figure 12. Box plots show that scores for Chinese students had the highest center (i.e., median math score) compared to the other subgroups at both 9th grade (*Med* = 61.9) and 11th grade (*Med* = 64.1), while Filipino students (*Med* = 53.7) had the lowest center in 9th grade and Southeast Asian students (*Med* = 53.1) had the lowest center at 11th grade. Other Asian students and South Asian students have a larger spread of scores compared to other subgroups at 9th grade, meaning scores had a wider range for these two groups than others. At 11th grade, the range of scores appeared to be relatively similar for all subgroups. Scores other Asian students, South Asian students, and Southeast Asian students appear to be left-skewed in 9th grade, with the second quartile having a wider range of scores than the third quartile. Eighth grade scores for Chinese students appear to be fairly symmetric with equivalent proportions of scores in the second and third quartile, while Filipino students appear to be right-skewed with scores covering more range between the third quartile than the second quartile. In 11th grade, scores appeared to be symmetric for Other Asian students, while scores were right-skewed for Southeast Asian students and Filipino students, and slightly left-skewed for Chinese students and South Asian students. Outliers were present at the low end of scores. In 9th grade, outliers were present for Chinese students and Southeast Asian students, while outliers were only present for Other Asian students at 11th grade.

ANOVAs were run to examine group differences in mean math scores across the original five Asian subgroup categories, as well as across the recoded three Asian subgroup categories. Descriptives of math scores across Asian subgroups are presented in

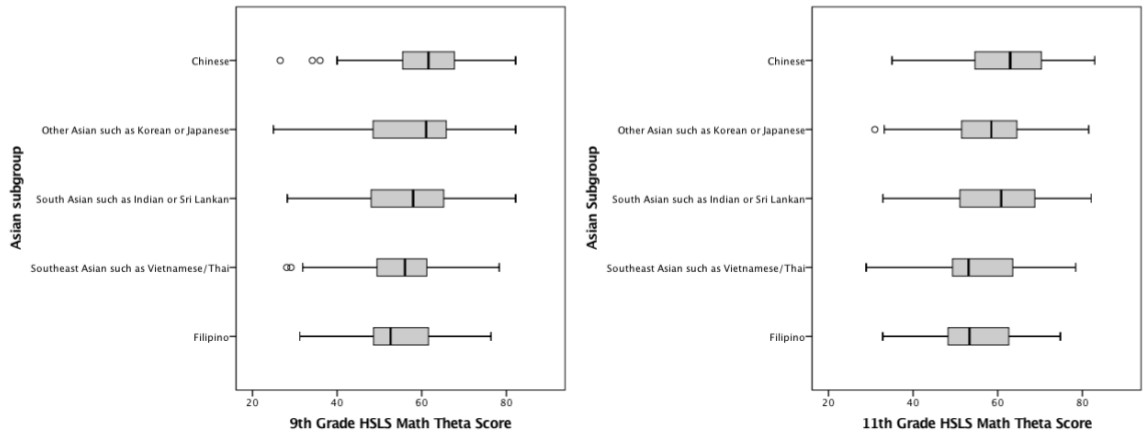


Figure 12. Box plots of HSLS math theta scores across Asian subgroups by grade.

Table 11, ANOVA model results are presented in Table 12, and Tukey HSD post-hoc tests are presented in Table 13 and 14. Across the original five Asian subgroup categories, there were statistically significant differences at the $\alpha = 0.010$ level only for 11th grade scores ($F = 4.600$, $p = 0.002$). Ninth grade scores were statistically significant at the $\alpha = 0.050$ level ($F = 2.973$, $p = 0.021$). Chinese students ($\bar{x} = 61.9$ and $SD = 8.8$ in 9th grade, $\bar{x} = 62.8$ and $SD = 9.6$ in 11th grade) typically had the highest math score at both grade levels. They had higher mean scores compared to Filipino students, but the differences were not significant at the $\alpha = 0.010$ level, a difference of 8.2 points in 9th grade and 8.6 points in 11th grade compared to Filipino students. Chinese students also had higher scores than Southeast Asian students, where scores were statistically significant at the $\alpha = 0.050$ level in 9th grade by 6.2 points and significant at the $\alpha = 0.010$ level in 11th grade by 8.1 points. Chinese students also had a higher mean math scores than South Asian students ($\bar{x} = 57.2$ and $SD = 11.4$ in 9th grade, $\bar{x} = 54.8$ and $SD = 9.7$ in 11th grade) and Other Asian students ($\bar{x} = 57.7$ and $SD = 12.0$ in 9th grade, $\bar{x} = 56.9$ and $SD = 11.2$ in 11th grade), but the differences were not statistically significant. There were no statistically significant differences in mean math scores between South Asian, Other Asian, Filipino, or Southeast Asian students at either grade level.

When the Asian subgroups were reclassified, East Asian students ($\bar{x}$ = 60.4 and *SD* = 10.1 in 9th grade, $\bar{x}$ = 60.8 and *SD* = 10.5 in 11th grade) had the highest mean math score at both grade levels, while Southeast Asian students had the lowest mean math scores ($\bar{x}$ = 55.0 and *SD* = 9.6 in 9th grade, $\bar{x}$ = 54.6 and *SD* = 9.3 in 11th grade). Differences between the mean scores of these two groups were statistically significant $\alpha = 0.010$ level by 5.4 points in 9th grade and 6.2 points in 11th grade. Mean scores for both of these groups were not significantly different from the mean math score of South Asian students ($\bar{x}$ = 57.2 and *SD* = 11.4 in 9th grade, $\bar{x}$ = 59.3 and *SD* = 11.3 in 11th grade).

Table 11
Descriptive statistics of HSLS math theta scores across Asian subgroups by grade.

	9th Grade				11th Grade			
	N	Mean	Median	StdDev	N	Mean	Median	StdDev
Original Asian subgroup categories								
Chinese	40	61.9	61.8	8.8	40	62.8	64.1	9.6
Other Asian	20	57.7	61.6	12.0	20	56.9	58.5	11.2
South Asian	40	57.2	57.9	11.4	40	59.3	61.7	11.3
Southeast Asian	50	55.7	56.3	9.7	50	54.8	53.1	9.7
Filipino	20	53.7	53.7	9.3	20	54.2	53.8	8.6
Recoded Asian subgroup categories								
East Asian	60	60.4	61.6	10.1	60	60.8	50.7	10.5
South Asian	40	57.2	57.9	11.4	40	59.3	53.1	11.3
Southeast Asian	70	55.0	54.6	9.6	80	54.6	62.9	9.3

Note. For the recoded Asian subgroup categories, Chinese and Other Asian students were combined, while Filipino and Southeast Asian students were combined. Sample sizes are rounded to the nearest tenth to protect the privacy of restricted-use data.

Table 12

One-way ANOVA results comparing mean HSLS math theta scores across Asian subgroups.

	Sum of Squares	Mean Square	F	*p*-value
Original Asian subgroup categories				
9th Grade [a]				
Between groups	1,230	300	2.973	0.021*
Within groups	16,520	100		
Total	17,750			
11th grade [b]				
Between groups	1,870	470	4.600	0.002**
Within groups	16,860	100		
Total	18,730			
Recoded Asian subgroup categories				
9th Grade [c]				
Between groups	936.210	470	4.510	0.012*
Within groups	16,810	100		
Total	17,750			
11th grade [d]				
Between groups	1,380	690	6.701	0.002**
Within groups	17,350	100		
Total	18,730			

Note. For the recoded Asian subgroup categories, Chinese and Other Asian students were combined into a category, while Filipino and Southeast Asian students were combined. The sum of squares and mean square values are rounded to the nearest tenth to protect the privacy of restricted-use data. Degrees of freedom are not reported to protect the privacy of restricted-use data.

[a] Levene's test of homogeneity (4, 160) = 1.158, *p*-value = 0.331.
[b] Levene's test of homogeneity (4, 170) = 1.048, *p*-value = 0.384.
[c] Levene's test of homogeneity (2, 160) = 0.886, *p*-value = 0.414.
[d] Levene's test of homogeneity (2, 170) = 1.872, *p*-value = 0.157.
[†] $p < .100$, *$p < .050$, **$p < .010$, ***$p < .001$

Table 13
Post-hoc comparisons of mean differences in HSLS math theta scores in 9th grade using Tukey HSD test.

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	Upper
Original Asian subgroup categories						
Chinese	Other Asian	4.1	2.8	0.572	-3.5	11.8
	South Asian	4.7	2.4	0.286	-1.9	11.3
	Southeast Asian	6.2*	2.2	0.044	0.1	12.2
	Filipino	8.2*	2.7	0.021	0.8	15.5
Other Asian	Chinese	-4.1	2.8	0.572	-11.8	3.5
	South Asian	0.6	2.8	1.000	-7.2	8.4
	Southeast Asian	2.0	2.7	0.941	-5.3	9.4
	Filipino	4.0	3.1	0.681	-4.4	12.5
South Asian	Chinese	-4.7	2.4	0.286	-11.3	1.9
	Other Asian	-0.6	2.8	1.000	-8.4	7.2
	Southeast Asian	1.5	2.3	0.966	-4.8	7.7
	Filipino	3.5	2.7	0.703	-4.0	11.0
Southeast Asian	Chinese	-6.2*	2.2	0.044	-12.2	-0.1
	Other Asian	-2.0	2.7	0.941	-9.4	5.3
	South Asian	-1.5	2.3	0.966	-7.7	4.8
	Filipino	2.0	2.6	0.935	-5.0	9.0
Filipino	Chinese	-8.2*	2.7	0.021	-15.5	-0.8
	Other Asian	-4.0	3.1	0.681	-12.5	4.4
	South Asian	-3.5	2.7	0.703	-11.0	4.0
	Southeast Asian	-2.0	2.6	0.935	-9.0	5.0
Recoded Asian subgroup categories						
East Asian	South Asian	3.2	2.2	0.300	-1.9	8.4
	Southeast Asian	5.4**	1.8	0.009	1.1	9.6
South Asian	East Asian	-3.2	2.2	0.300	-8.4	1.9
	Southeast Asian	2.1	2.1	0.570	-2.8	7.1
Southeast Asian	East Asian	-5.4**	1.8	0.009	-9.6	-1.1
	South Asian	-2.1	2.1	0.570	-7.1	2.8

Note. For the recoded Asian subgroup categories, Chinese and Other Asian students were combined, while Filipino and Southeast Asian students were combined.
$^{\dagger}p < .100$, $^{*}p < .050$, $^{**}p < .010$, $^{***}p < .001$

Table 14
Post-hoc comparisons of mean differences in HSLS math theta scores in 11th grade using Tukey HSD test.

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	95% Confidence Interval Upper
Original Asian subgroup categories						
Chinese	Other Asian	6.0	2.7	0.194	-1.6	13.5
	South Asian	3.5	2.3	0.571	-3.0	9.9
	Southeast Asian	8.1**	2.1	0.002	2.2	13.9
	Filipino	8.6*	2.6	0.011	1.3	15.8
Other Asian	Chinese	-6.0	2.7	0.194	-13.5	1.6
	South Asian	-2.5	2.8	0.898	-10.2	5.2
	Southeast Asian	2.1	2.6	0.930	-5.1	9.3
	Filipino	2.6	3.0	0.909	-5.8	11.0
South Asian	Chinese	-3.5	2.3	0.571	-9.9	3.0
	Other Asian	2.5	2.8	0.898	-5.2	10.2
	Southeast Asian	4.6	2.2	0.227	-1.5	10.6
	Filipino	5.1	2.7	0.315	-2.3	12.5
Southeast Asian	Chinese	-8.1**	2.1	0.002	-13.9	-2.2
	Other Asian	-2.1	2.6	0.930	-9.3	5.1
	South Asian	-4.6	2.2	0.227	-10.6	1.5
	Filipino	0.5	2.5	1.000	-6.4	7.4
Filipino	Chinese	-8.6*	2.6	0.011	-15.8	-1.3
	Other Asian	-2.6	3.0	0.909	-11.0	5.8
	South Asian	-5.1	2.7	0.315	-12.5	2.3
	Southeast Asian	-0.5	2.5	1.000	-7.4	6.4
Recoded Asian subgroup categories						
East Asian	South Asian	1.4	2.1	0.788	-3.7	6.5
	Southeast Asian	6.2**	1.8	0.002	2.0	10.3
South Asian	East Asian	-1.4	2.1	0.788	-6.5	3.7
	Southeast Asian	4.8	2.1	0.058	-0.1	9.6
Southeast Asian	East Asian	-6.2**	1.8	0.002	-10.3	-2.0
	South Asian	-4.8	2.1	0.058	-9.6	0.1

Note. For the recoded Asian subgroup categories, Chinese and Other Asian students were combined, while Filipino and Southeast Asian students were combined.
$^{\dagger}p < .100$, $^{*}p < .050$, $^{**}p < .010$, $^{***}p < .001$

MCA. Figure 13 displays the box plots of math scores across Asian subgroups. Students with a Chinese-Mandarin home language had the highest center (i.e., median score) at 59.0 in 8th grade and 59.3 at 11th grade. Students with a Lao home language had the lowest center of 46.6 in 8th grade, while students with Other Southeast Asian language had the lowest center at 38.3. Students with an Other Southeast Asian language also had the largest spread of scores at both grade levels, while students with Hmong, Vietnamese, or Lao home language had scores closer together at 11th grade. Box plots of 8th grade scores for students with a Hmong, Vietnamese, or South Asian home language appeared to be symmetric, with scores spread fairly evenly between the second and third quartiles and with the first and fourth quartiles having roughly an equivalent range of scores. For students with an Other East Asian language, English, Khmer, or Other Southeast Asian language, box plots appeared to be somewhat left-skewed with more scores spread across quartiles 1 and 2 than quartiles 3 and 4. Conversely, scores for students with a Chinese-Mandarin, Other language, or Lao home language were slightly right skewed with scores more spread out across quartile 3 than 2. As for 11th grade scores, box plots appeared to be symmetric for English, Khmer, Lao, Hmong and Other language groups. Box plots for those with a South Asian, Other East Asian, or Vietnamese home language were left-skewed, while scores for students with a Chinese-Mandarin or Other Southeast Asian home language were right-skewed. Outliers were present at the low and high end of scores for students with a home language of English, Vietnamese, Other language, or Hmong in 8th grade, whereas, there was a presence of outliers on the low end for students with an Other East Asian language and on the high end for students with a Lao home language. In 11th grade, outliers were present at the low end of scores for the following

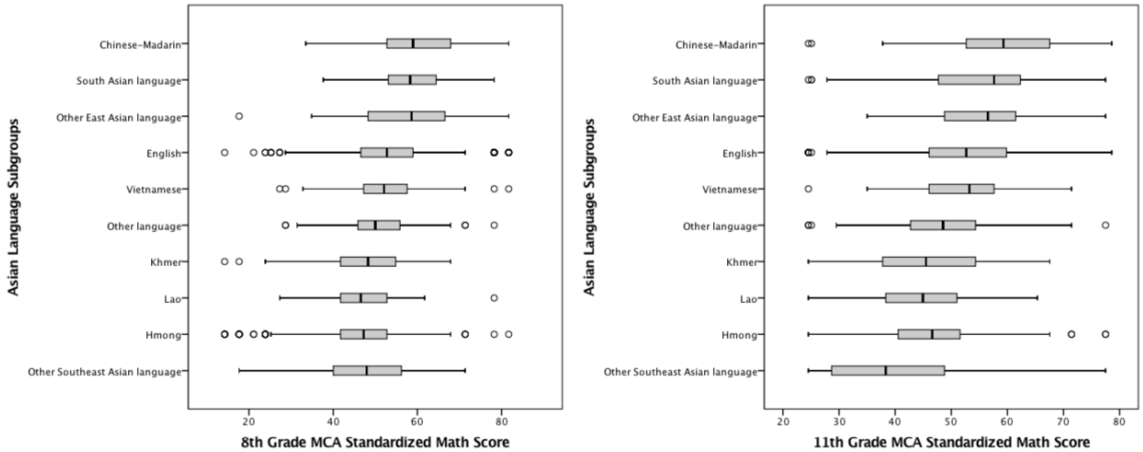


Figure 13. Box plots of MCA standardized math scores across Asian subgroups by grade.

language subgroups: Chinese-Mandarin, South Asian language, English, and Vietnamese. Additionally, there were low and high outliers among students with an Other language, and the Hmong language subgroup also had outliers on the high end of scores.

Descriptives of math scores across Asian subgroups are presented in Table 15, ANOVA model results are presented in Table 16, and Tukey HSD post-hoc tests are presented in Table 17 and 18. ANOVA results showed there were statistically significant differences in mean math scores across Asian language subgroups in both 8th grade (F(9, 2,746) = 51.704, p = 0.000) and 11th grade (F(9, 3,289) = 83.382, p = 0.000) (Table X). At both grade levels, Chinese-Mandarin students ($\bar{x}$ = 60.1 and SD = 10.5 in 8th grade, $\bar{x}$ = 59.2 and SD = 10.2 in 11th grade), Other East Asian language students ($\bar{x}$ = 56.6 and SD = 9.5 in 8th grade, $\bar{x}$ = 56.3 and SD = 10.0 in 11th grade), and South Asian language students ($\bar{x}$ = 57.8 and SD = 9.5 in 8th grade, $\bar{x}$ = 48.7 and SD = 11.9 in 11th grade) had the highest mean math scores. Differences in mean math scores across these three groups were not significantly different, with the exception of Chinese-Mandarin students having a significantly higher mean score than South Asian language students in 11th grade by 4.7 points.

Compared to the Other language groups, students with a Chinese-Mandarin home language had a significantly higher mean math score at both grade levels than those with an English home language ($\bar{x}$ = 53.2 and *SD* = 10.2 in both 8th grade and 11th grade), Vietnamese home language ($\bar{x}$ = 52.3 and *SD* = 8.5, $\bar{x}$ = 52.5 and *SD* = 8.5 in 11th grade), Other language ($\bar{x}$ = 50.6 and *SD* = 11.0, $\bar{x}$ = 48.7 and *SD* = 11.9 in 11th grade), Khmer home language ($\bar{x}$ = 47.4 and *SD* = 11.0, $\bar{x}$ = 45.4 and *SD* = 10.8 in 11th grade), Lao home language ($\bar{x}$ = 46.6 and *SD* = 8.1, $\bar{x}$ = 44.4 and *SD* = 8.7 in 11th grade), Hmong home language ($\bar{x}$ = 46.5 and *SD* = 9.0, $\bar{x}$ = 46.0 and *SD* = 8.6 in 11th grade), and Other Southeast Asian language ($\bar{x}$ = 46.3 and *SD* = 13.4, $\bar{x}$ = 48.7 and *SD* = 11.9 in 11th grade). Differences ranged between 6.9 to 13.9 points in 8th grade and 6.0 to 19.7 points in 11th grade. The South Asian language student group also had a statistically significant higher mean math scores than students with the following home languages: English, Vietnamese, Other language, Khmer, Lao, Hmong, and Other Southeast Asian language in 8th grade with mean differences ranging from 4.6 to 11.5 points. In 11th grade, South Asian language students continued to have a higher mean math score than these groups, but differences were not significant when compared to mean math scores of students with an English and Vietnamese home language. Mean differences between the South Asian language student group and the other five groups ranged between 5.8 and 15.0 points. The Other East Asian language student group had a higher mean math score at both grade levels than Khmer, Hmong, Lao, and Other Southeast Asian language student groups—with the addition of the Other Asian language student group in 11th grade. Mean differences in math scores ranged from 9.2 to 10.3 points in 8th grade and 7.6 to 16.9 points in 11th grade.

Students with an English home language and Vietnamese home language had mean math scores in the middle relative to the other Asian language groups. Mean math scores did not significantly differ between these two groups at both grade levels, but they did have significantly lower mean scores compared to students with a Chinese-Mandarin home language at both grade levels and students with a South Asian home language in 8th grade. In contrast, the two groups had significantly higher mean math scores than Khmer, Hmong, Lao, and Other Southeast Asian language student groups at both grade levels. Students with an English home language also had a significantly higher score than Other language students in 8th grade.

The Other language student group scored in the low middle range compared to the Other language groups. In 8th grade, the mean math score for this group was significantly lower from Chinese-Mandarin students by 9.5 points and South Asian language students by 7.2 points, but not significantly different from any of the other groups. In 11th grade, the Other language student group's mean math score was significantly lower compared to Chinese-Mandarin students by 10.5 points, Other East Asian language students by 7.6 points, South Asian language by 5.8 points, and English students by 4.5 points. Additionally, this group had a significantly higher mean score than the Other Southeast Asian language group by 9.3 points.

Khmer, Hmong, Lao, and Other Southeast Asian language groups had the lowest mean math scores at both grade levels. Mean math scores in 8th grade were not significantly different between paired combinations of these four groups. However, in

11th grade, the Other Southeast Asian language group had a significantly lower mean math score than all other language groups, including Hmong students by 6.5, Khmer students by 6 points, and Lao students by 4.9 points. Hmong students, Khmer students, and Lao students did not have significantly different mean math scores at 11th grade from each other, but continued to have significantly lower mean math scores compared to Chinese-Mandarin, Other East Asian language, South Asian language, English, and Vietnamese language groups.

Table 15

Descriptives of MCA standardized math scores across Asian subgroups by grade.

Asian language groups	8th Grade				11th Grade			
	N	Mean	Median	StdDev	N	Mean	Median	StdDev
Chinese-Mandarin	107	60.1	59.0	10.5	125	59.2	59.3	10.2
South Asian language	72	57.8	58.3	9.5	87	54.5	57.6	12.3
Other East Asian language	40	56.6	58.6	12.5	55	56.3	56.5	10.0
English	938	53.2	52.8	10.2	938	53.2	52.7	10.2
Vietnamese	198	52.3	52.1	8.5	242	52.5	53.2	8.5
Other language	48	50.6	50.0	11.0	54	48.7	48.5	11.9
Khmer	78	47.4	48.3	11.0	107	45.4	45.5	10.8
Lao	93	46.6	46.6	8.1	112	44.4	44.9	8.7
Hmong	1,119	46.5	47.3	9.0	1,410	46.0	46.6	8.6
Other Southeast Asian language	63	46.3	47.9	13.4	160	39.5	38.3	12.4

Table 16

One-way ANOVA results comparing mean MCA standardized math scores across Asian language subgroups by grade.

	Sum of Squares	df	Mean Square	F	*p*-value
8th Grade [a]					
Between groups	43,804.286	9	4,867.143	51.704	0.000***
Within groups	258,492.030	2,746	94.134		
Total	302,296.316	2,755			
11th Grade [b]					
Between groups	68,925.701	9	7,658.411	83.382	0.000***
Within groups	301,259.311	3,280	91.847		
Total	370,185.012	3,289			

[a] Levene's test of homogeneity (9, 2,746) = 5.471, *p*-value = 0.000***.

[b] Levene's test of homogeneity (9, 3,280) = 12.545, *p*-value = 0.000***.

†*p* < .100, **p* < .050, ***p* < .010, ****p* < .001

Table 17

Post-hoc comparisons of mean differences in MCA standardized math scores in 8th grade using Tukey HSD test.

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	Upper
Chinese-Mandarin	South Asian language	2.3	1.5	0.857	-2.3	7.0
	Other East Asian language	3.6	1.8	0.611	-2.1	9.3
	English	6.9***	1.0	0.000	3.8	10.0
	Vietnamese	7.8***	1.2	0.000	4.1	11.5
	Other language	9.5***	1.7	0.000	4.2	14.8
	Khmer	12.8***	1.4	0.000	8.2	17.4
	Lao	13.5***	1.4	0.000	9.2	17.9
	Hmong	13.6***	1.0	0.000	10.5	16.7
	Other Southeast Asian language	13.9***	1.5	0.000	9.0	18.8
South Asian language	Chinese-Mandarin	-2.3	1.5	0.857	-7.0	2.3
	Other East Asian language	1.2	1.9	1.000	-4.8	7.3
	English	4.6**	1.2	0.005	0.8	8.3
	Vietnamese	5.5**	1.3	0.002	1.3	9.7
	Other language	7.2**	1.8	0.003	1.4	12.9
	Khmer	10.4***	1.6	0.000	5.4	15.5
	Lao	11.2***	1.5	0.000	6.4	16.0
	Hmong	11.3***	1.2	0.000	7.5	15.0
	Other Southeast Asian language	11.5***	1.7	0.000	6.3	16.9
Other East Asian language	Chinese-Mandarin	-3.6	1.8	0.611	-9.3	2.1
	South Asian language	-1.2	1.9	1.000	-7.3	4.8
	English	3.3	1.6	0.508	-1.6	8.3
	Vietnamese	4.3	1.7	0.251	-1.1	9.6
	Other language	5.9	2.1	0.119	-0.7	12.5
	Khmer	9.2***	1.9	0.000	3.2	15.2
	Lao	10.0***	1.8	0.000	4.2	15.8
	Hmong	10.0***	1.6	0.000	5.1	15.0
	Other Southeast Asian language	10.3***	2.0	0.000	4.1	16.5

$^{\dagger}p < .100$, $*p < .050$, $**p < .010$, $***p < .001$

Table 17

Post-hoc comparisons of mean differences in MCA standardized math scores in 8th grade using Tukey HSD test (continued…).

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	95% Confidence Interval Upper
English	Chinese-Mandarin	-6.9***	1.0	0.000	-10.0	-3.8
	South Asian language	-4.6**	1.2	0.005	-8.3	-0.8
	Other East Asian language	-3.3	1.6	0.508	-8.3	1.6
	Vietnamese	0.9	0.8	0.969	-1.5	3.3
	Other language	2.6	1.4	0.730	-2.0	7.1
	Khmer	5.9***	1.1	0.000	2.3	9.5
	Lao	6.6***	1.1	0.000	3.3	10.0
	Hmong	6.7***	0.4	0.000	5.4	8.1
	Other Southeast Asian language	7.0***	1.3	0.000	3.0	11.0
Vietnamese	Chinese-Mandarin	-7.8***	1.2	0.000	-11.5	-4.1
	South Asian language	-5.5**	1.3	0.002	-9.7	-1.3
	Other East Asian language	-4.3	1.7	0.251	-9.6	1.1
	English	-0.9	0.8	0.969	-3.3	1.5
	Other language	1.7	1.6	0.988	-3.3	6.6
	Khmer	5.0**	1.3	0.005	0.9	9.1
	Lao	5.7***	1.2	0.000	1.9	9.6
	Hmong	5.8***	0.7	0.000	3.4	8.2
	Other Southeast Asian language	6.1**	1.4	0.001	1.6	10.5
Other language	Chinese-Mandarin	-9.5***	1.7	0.000	-14.8	-4.2
	South Asian language	-7.2**	1.8	0.003	-12.9	-1.4
	Other East Asian language	-5.9	2.1	0.119	-12.5	0.7
	English	-2.6	1.4	0.730	-7.1	2.0
	Vietnamese	-1.7	1.6	0.988	-6.6	3.3
	Khmer	3.3	1.8	0.704	-2.4	8.9
	Lao	4.0	1.7	0.359	-1.4	9.5
	Hmong	4.1	1.4	0.112	-0.4	8.7
	Other Southeast Asian language	4.4	1.9	0.349	-1.5	10.3

†*p* < .100, **p* < .050, ***p* < .010, ****p* < .001

Table 17

Post-hoc comparisons of mean differences in MCA standardized math scores in 8th grade using Tukey HSD test (continued…).

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	Upper
Khmer	Chinese-Mandarin	-12.8***	1.4	0.000	-17.4	-8.2
	South Asian language	-10.4***	1.6	0.000	-15.5	-5.4
	Other East Asian language	-9.2***	1.9	0.000	-15.2	-3.2
	English	-5.9***	1.1	0.000	-9.5	-2.3
	Vietnamese	-5.0**	1.3	0.005	-9.1	-0.9
	Other language	-3.3	1.8	0.704	-8.9	2.4
	Lao	0.8	1.5	1.000	-4.0	5.5
	Hmong	0.8	1.1	0.999	-2.8	4.4
	Other Southeast Asian language	1.1	1.6	1.000	-4.1	6.3
Lao	Chinese-Mandarin	-13.5***	1.4	0.000	-17.9	-9.2
	South Asian language	-11.2**	1.5	0.000	-16.0	-6.4
	Other East Asian language	-10.0***	1.8	0.000	-15.8	-4.2
	English	-6.6***	1.1	0.000	-10.0	-3.3
	Vietnamese	-5.7***	1.2	0.000	-9.6	-1.9
	Other language	-4.0	1.7	0.359	-9.5	1.4
	Khmer	-0.8	1.5	1.000	-5.5	4.0
	Hmong	0.1	1.0	1.000	-3.2	3.4
	Other Southeast Asian language	0.3	1.6	1.000	-4.7	5.4
Hmong	Chinese-Mandarin	-13.6***	1.0	0.000	-16.7	-10.5
	South Asian language	-11.3***	1.2	0.000	-15.0	-7.5
	Other East Asian language	-10.0***	1.6	0.000	-15.0	-5.1
	English	-6.7***	0.4	0.000	-8.1	-5.4
	Vietnamese	-5.8***	0.7	0.000	-8.2	-3.4
	Other language	-4.1	1.4	0.112	-8.7	0.4
	Khmer	-0.8	1.1	0.999	-4.4	2.8
	Lao	-0.1	1.0	1.000	-3.4	3.2
	Other Southeast Asian language	0.3	1.3	1.000	-3.7	4.3

$^{\dagger}p < .100$, $*p < .050$, $**p < .010$, $***p < .001$

Table 17

Post-hoc comparisons of mean differences in MCA standardized math scores 8th grade using Tukey HSD test (continued…).

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	95% Confidence Interval Upper
Other Southeast Asian language	Chinese-Mandarin	-13.9***	1.5	0.000	-18.8	-9.0
	South Asian language	-11.5***	1.7	0.000	-16.9	-6.3
	Other East Asian language	-10.3***	2.0	0.000	-16.5	-4.1
	English	-7.0***	1.3	0.000	-11.0	-3.0
	Vietnamese	-6.1**	1.4	0.001	-10.5	-1.6
	Other language	-4.4	1.9	0.349	-10.3	1.5
	Khmer	-1.1	1.6	1.000	-6.3	4.1
	Lao	-0.3	1.6	1.000	-5.4	4.7
	Hmong	-0.3	1.3	1.000	-4.3	3.7

†$p < .100$, *$p < .050$, **$p < .010$, ***$p < .001$

Table 18

Post-hoc comparisons of mean differences in MCA standardized math scores in 11th grade using Tukey HSD test.

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	95% Confidence Interval Upper
Chinese-Mandarin	South Asian language	4.7*	1.338	0.016	0.5	8.9
	Other East Asian language	2.9	1.551	0.704	-2.0	7.8
	English	6.0***	0.913	0.000	3.1	8.9
	Vietnamese	6.7***	1.056	0.000	3.3	10.0
	Other language	10.5***	1.561	0.000	5.6	15.4
	Khmer	13.8***	1.262	0.000	9.8	17.8
	Lao	14.8***	1.247	0.000	10.9	18.8
	Hmong	13.2***	0.894	0.000	10.4	16.0
	Other Southeast Asian language	19.7***	1.144	0.000	16.1	23.3
South Asian language	Chinese-Mandarin	-4.7*	1.338	0.016	-8.9	-0.5
	Other East Asian language	-1.8	1.651	0.983	-7.1	3.4
	English	1.3	1.074	0.969	-2.1	4.7
	Vietnamese	1.9	1.198	0.838	-1.9	5.7
	Other language	5.8*	1.660	0.018	0.5	11.0
	Khmer	9.0***	1.384	0.000	4.7	13.4
	Lao	10.1***	1.370	0.000	5.8	14.5
	Hmong	8.5***	1.059	0.000	5.1	11.8
	Other Southeast Asian language	15.0***	1.277	0.000	11.0	19.1
Other East Asian language	Chinese-Mandarin	-2.9	1.551	0.704	-7.8	2.0
	South Asian language	1.8	1.651	0.983	-3.4	7.1
	English	3.2	1.330	0.343	-1.1	7.4
	Vietnamese	3.8	1.432	0.198	-0.8	8.3
	Other language	7.6**	1.836	0.001	1.8	13.4
	Khmer	10.9***	1.590	0.000	5.9	15.9
	Lao	12.0***	1.578	0.000	7.0	17.0
	Hmong	10.3***	1.317	0.000	6.2	14.5
	Other Southeast Asian language	16.9***	1.498	0.000	12.1	21.6

$^{\dagger}p < .100$, $^{*}p < .050$, $^{**}p < .010$, $^{***}p < .001$

Table 18

Post-hoc comparisons of mean differences in MCA standardized math scores in 11th grade using Tukey HSD test (continued…).

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	95% Confidence Interval Upper
English	Chinese-Mandarin	-6.0***	0.913	0.000	-8.9	-3.1
	South Asian language	-1.3	1.074	0.969	-4.7	2.1
	Other East Asian language	-3.2	1.330	0.343	-7.4	1.1
	Vietnamese	0.6	0.691	0.996	-1.6	2.8
	Other language	4.5*	1.341	0.030	0.2	8.7
	Khmer	7.7***	0.978	0.000	4.6	10.8
	Lao	8.8***	0.958	0.000	5.8	11.8
	Hmong	7.2***	0.404	0.000	5.9	8.5
	Other Southeast Asian language	13.7***	0.820	0.000	11.1	16.3
Vietnamese	Chinese-Mandarin	-6.7***	1.056	0.000	-10.0	-3.3
	South Asian language	-1.9	1.198	0.838	-5.7	1.9
	Other East Asian language	-3.8	1.432	0.198	-8.3	0.8
	English	-0.6	0.691	0.996	-2.8	1.6
	Other language	3.8	1.442	0.190	-0.7	8.4
	Khmer	7.1***	1.113	0.000	3.6	10.6
	Lao	8.2***	1.095	0.000	4.7	11.6
	Hmong	6.6***	0.667	0.000	4.4	8.7
	Other Southeast Asian language	13.1***	0.977	0.000	10.0	16.2
Other language	Chinese-Mandarin	-10.5***	1.561	0.000	-15.4	-5.6
	South Asian language	-5.8*	1.660	0.018	-11.0	-0.5
	Other East Asian language	-7.6**	1.836	0.001	-13.4	-1.8
	English	-4.5*	1.341	0.030	-8.7	-0.2
	Vietnamese	-3.8	1.442	0.190	-8.4	0.7
	Khmer	3.3	1.600	0.572	-1.8	8.3
	Lao	4.3	1.588	0.162	-0.7	9.4
	Hmong	2.7	1.329	0.571	-1.5	6.9
	Other Southeast Asian language	9.3***	1.508	0.000	4.5	14.0

†$p < .100$, *$p < .050$, **$p < .010$, ***$p < .001$

Table 18

Post-hoc comparisons of mean differences in MCA standardized math scores in 11th grade using Tukey HSD test (continued…).

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval Lower	Upper
Khmer	Chinese-Mandarin	-13.8***	1.262	0.000	-17.8	-9.8
	South Asian language	-9.0***	1.384	0.000	-13.4	-4.7
	Other East Asian language	-10.9***	1.590	0.000	-15.9	-5.9
	English	-7.7***	0.978	0.000	-10.8	-4.6
	Vietnamese	-7.1***	1.113	0.000	-10.6	-3.6
	Other language	-3.3	1.600	0.572	-8.3	1.8
	Lao	1.1	1.296	0.998	-3.0	5.2
	Hmong	-0.6	0.961	1.000	-3.6	2.5
	Other Southeast Asian language	6.0***	1.197	0.000	2.2	9.8
Lao	Chinese-Mandarin	-14.8***	1.247	0.000	-18.8	-10.9
	South Asian language	-10.1***	1.370	0.000	-14.5	-5.8
	Other East Asian language	-12.0***	1.578	0.000	-17.0	-7.0
	English	-8.8***	0.958	0.000	-11.8	-5.8
	Vietnamese	-8.2***	1.095	0.000	-11.6	-4.7
	Other language	-4.3	1.588	0.162	-9.4	0.7
	Khmer	-1.1	1.296	0.998	-5.2	3.0
	Hmong	-1.6	0.941	0.779	-4.6	1.4
	Other Southeast Asian language	4.9**	1.181	0.001	1.2	8.6
Hmong	Chinese-Mandarin	-13.2***	0.894	0.000	-16.0	-10.4
	South Asian language	-8.5***	1.059	0.000	-11.8	-5.1
	Other East Asian language	-10.3***	1.317	0.000	-14.5	-6.2
	English	-7.2***	0.404	0.000	-8.5	-5.9
	Vietnamese	-6.6***	0.667	0.000	-8.7	-4.4
	Other language	-2.7	1.329	0.571	-6.9	1.5
	Khmer	0.6	0.961	1.000	-2.5	3.6
	Lao	1.6	0.941	0.779	-1.4	4.6
	Other Southeast Asian language	6.5***	0.799	0.000	4.0	9.1

$^{\dagger}p < .100$, $^{*}p < .050$, $^{**}p < .010$, $^{***}p < .001$

Table 18

Post-hoc comparisons of mean differences in MCA standardized math scores in 11th grade using Tukey HSD test (continued…).

Comparisons		Mean difference	Std Error	*p*-value	95% Confidence Interval	
					Lower	Upper
Other Southeast Asian language	Chinese-Mandarin	-19.7***	1.144	0.000	-23.3	-16.1
	South Asian language	-15.0***	1.277	0.000	-19.1	-11.0
	Other East Asian language	-16.9***	1.498	0.000	-21.6	-12.1
	English	-13.7***	0.820	0.000	-16.3	-11.1
	Vietnamese	-13.1***	0.977	0.000	-16.2	-10.0
	Other language	-9.2***	1.508	0.000	-14.0	-4.5
	Khmer	-6.0***	1.197	0.000	-9.8	-2.2
	Lao	-4.9**	1.181	0.001	-8.6	-1.2
	Hmong	-6.5***	0.799	0.000	-9.1	-4.0

$^{\dagger}p < .100$, $*p < .050$, $**p < .010$, $***p < .001$

Comparison of Math Achievement across Racial groups with the Use of an Aggregated Asian group Versus Disaggregated Asian Subgroups

HSLS. The final models predicting math achievement when using an aggregate Asian group and disaggregated Asian subgroups are presented in Equations 3 to 6.

Use of an aggregate Asian group

Model A.1: White students as the reference group

$$\begin{aligned}
Math\ Achievement &= 54.49 - 0.49(Grade) + 2.21(Asian) - 4.60(Black) \\
&- 2.12(PovertyThreshold) - 1.86(FirstGenCollege) \\
&- 7.57(PreAorAI) - 4.96(Geometry) + 2.39(Trig) \\
&+ 7.56(PreCalc) + 9.54(Calc) + 4.55(Stat) \\
&- 3.03(OtherMath) - 3.11(NoAdvMath)
\end{aligned} \tag{3}$$

Model A.2: Black students as the reference group

$$\begin{aligned}
Math\ Achievement &= 49.89 - 0.49(Grade) + 6.81(Asian) + 4.60(White) \\
&- 2.12(PovertyThreshold) - 1.86(FirstGenCollege) \\
&- 7.57(PreAorAI) - 4.96(Geometry) + 2.39(Trig) \\
&+ 7.56(PreCalc) + 9.54(Calc) + 4.55(Stat) \\
&- 3.03(OtherMath) - 3.11(NoAdvMath)
\end{aligned} \tag{4}$$

Use of disaggregated Asian subgroups

Model B.1: White students as the reference group

$$\begin{aligned}
Math\ Achievement &= 54.51 - 0.48(Grade) + 5.52(EastAsian) \\
&- 1.046(SouthAsian) + 0.71(SoutheastAsian) - 4.61(Black) \\
&- 2.08(PovertyThreshold) - 1.86(FirstGenCollege) \\
&- 7.45(PreAorAI) - 4.91(Geometry) + 2.39(Trig) \\
&+ 7.56(PreCalc) + 9.62(Calc) + 4.55(Stat) \\
&- 3.06(OtherMath) - 3.16(NoAdvMath)
\end{aligned} \tag{5}$$

Model B.2: Black students as the reference group

$$\begin{aligned}
Math\ Achievement &= 49.91 - 0.48(Grade) + 10.13(EastAsian) \\
&- 3.56(SouthAsian) + 5.32(SoutheastAsian) - 4.61(White) \\
&- 2.08(PovertyThreshold) - 1.86(FirstGenCollege) \\
&- 7.45(PreAorAI) - 4.91(Geometry) + 2.39(Trig) \\
&+ 7.56(PreCalc) + 9.62(Calc) + 4.55(Stat) \\
&- 3.06(OtherMath) - 3.16(NoAdvMath)
\end{aligned} \tag{6}$$

F-tests of the fixed effects are provided in Table 19 for both models, while Table 20 displays the parameter estimates of fixed effects, variance of random effects, number of estimated parameters, and model fit statistics. Fitted line graphs comparing the change in math achievement between grade levels across race while controlling for covariates are displayed in Figures 14 and 15. Five predictors significantly contributed to predicting math achievement, including race, 185% poverty threshold, first-generation college student status, type of math course, and taking advanced math coursework. Between 9^{th} and 11^{th} grade, the estimated math achievement increases by about 0.48 points on average in both models. The fixed effect for this change is statistically significantly different from zero at the $\alpha = 0.050$ level. LME models do not have an equivalent R^2 value (i.e., the proportion of total variance explained by the model or how closely observed values reflect fitted values) like linear fixed effects models do, given their complexities of having variance associated with both the random and fixed effects. However, an Intraclass Correlation Coefficient (ICC) can be computed for LME models to assess how much of the variance is accounted for by the random effects (West, Welch, & Galecki, 2015). In both Models A and B, allowing intercepts to vary accounted for about 58% of the variance of the random effects.

Math achievement was significantly different across race while accounting for the covariates. Model A showed that Asian students, as an aggregate group, had a higher mean math achievement by a difference of 2.21 points compared to White students while holding all else constant. However, this average difference was only statistically significant at the $\alpha = 0.050$ level. In comparison with the use of disaggregated Asian subgroups, Model B showed that only East Asian students had significantly higher math achievement on average by 5.52 points while accounting for all covariates. Southeast Asian students had slightly higher mean math scores than White students by 0.71 points on average while controlling for all covariates, whereas South Asian students had lower scores by 1.05 points on average. These differences were not statistically different. Asian students, when grouped aggregately and disaggregately, had higher mean math scores than Black students. In model A, Asian students had a higher mean math score than Black students by 6.81 points while holding all other covariates constant. In model B with all covariates accounted for, East Asian students and Southeast Asian students had a significantly higher mean math score by 10.13 points and 5.32 points, respectively, compared to Black students. South Asian students also had a higher mean math score by 3.56 compared to Black students after accounting for covariates, but the difference was not statistically significant. Math achievement typically decreased if students were from families with household incomes below the 185% U.S. poverty threshold (β = -2.12 in Model A and β = -2.08 in Model B), will be first-generation college students (β = -1.86 in both models), and/or never took advanced math courses in high school (β = -3.11 in Model A and β = -3.16 in Model B). Across types of math courses taken, students who took trigonometry (β = 2.39 in both models), pre-calculus (β = 7.56 in both models), calculus (β = 9.54 in Model A and β = 9.62 in Model B), and statistics (β = 4.55 in both models) typically had higher mean math achievement than those who took algebra II/III. Students who took pre-algebra/algebra I (β = -7.57 in Model A and β = -7.45 in Model B), geometry (β = -4.96 in Model A and β = -4.91 in Model B), or some other math course (β = -3.03 in Model A and β = -3.06 in Model B) had typically lower math achievement than those who took algebra II/III.

Table 19
Type III Tests of fixed effects predicting HSLS math theta scores.

	F	*p*-value
Model A: Aggregated Asian group		
Intercept	14,020.858	0.000
Grade	6.158	0.013
Race: Aggregate Asian	45.595	0.000
Below 185% poverty threshold	21.993	0.000
First-generation college student	15.757	0.000
Most advanced math course taken	67.089	0.000
Never taken advanced math course	27.241	0.000
Model B: Disaggregated Asian subgroup		
Intercept	5,911.927	0.000
Grade	6.111	0.014
Race: Disaggregate Asian	24.873	0.000
Below 185% poverty threshold	21.256	0.000
First-generation college student	15.854	0.000
Most advanced math course taken	66.853	0.000
Never taken advanced math course	28.308	0.000

Notes. Degrees of freedom are not reported to protect the privacy of restricted-use data.

Table 20

Comparison of parameter estimates, number of parameters estimated, error variance, and model fit indices predicting HSLS math theta scores between 9th grade and 11th grade, with Models A.1 and A.2 using an aggregated Asian group and Models B.1 and B.2 using disaggregated Asian groups.

	Aggregated Asian Group		Disaggregated Asian Groups	
	Model A.1 (Race reference group = White)	Model A.2 (Race reference group = Black)	Model B.1 (Race reference group = White)	Model B.2 (Race reference group = Black)
Fixed Effects				
Intercept	54.487***	49.891***	54.513***	49.906***
Slope	0.485*	0.485*	0.484*	0.484*
Asian	2.211*	6.807***		
East Asian			5.521***	10.127***
South Asian			-1.046	3.561
Southeast Asian			0.713	5.320***
Black	-4.595***		-4.607***	
White		4.595***		4.607***
Below 185% poverty threshold	-2.120***	-2.120***	-2.081***	-2.081***
First-generation college student	-1.861***	-1.861***	-1.863***	-1.863***
Pre-Algebra/Algebra I	-7.565***	-7.565***	-7.453***	-7.453***
Geometry	-4.961***	-4.961***	-4.913***	-4.913***
Trigonometry	2.385**	2.385**	2.393**	2.393**
Pre-Calculus	7.562***	7.562***	7.557***	7.557***
Calculus	9.543***	9.543***	9.615***	9.615***
Statistics	4.553***	4.553***	4.554***	4.554***
Other Math Course	-3.030***	-3.030***	-3.063***	-3.063***
Never taken advanced math course	-3.111***	-3.111***	-3.163***	-3.163***

*$p < .050$, **$p < .010$, ***$p < .000$

Table 20

Comparison of parameter estimates, number of parameters estimated, error variance, and model fit indices predicting HSLS math theta scores between 9th grade and 11th grade, with Models A.1 and A.2 using an aggregated Asian group and Models B.1 and B.2 using disaggregated Asian groups (continued…).

	Aggregated Asian Group		Disaggregated Asian Groups	
	Model A.1 (Race reference group = White)	Model A.2 (Race reference group = Black)	Model B.1 (Race reference group = White)	Model B.2 (Race reference group = Black)
Variance of Random Effects				
Intercepts	33.181	33.181	32.887	32.887
Error	23.942	23.942	23.963	23.963
Number of estimated parameters	16	16	18	18
Model fit statistics				
AIC	16,741.541	16,741.541	16,725.831	16,725.831
BIC	16,834.803	16,834.803	16,830.735	16,830.735

**p* < .050, ***p* < .010, ****p* < .000

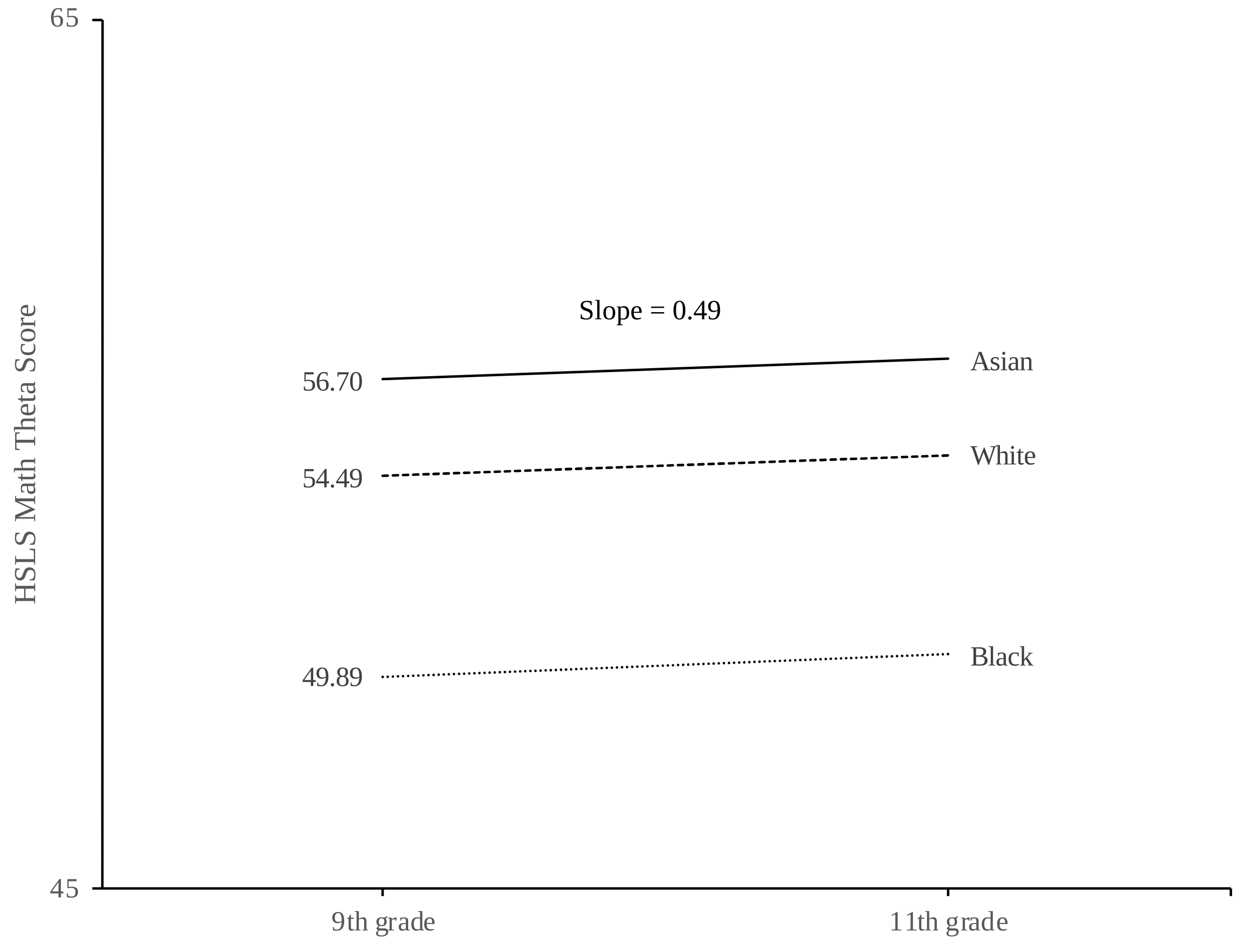


Figure 14. Fitted regression line of the change in HSLS math theta scores from 9th grade to 11th grade by race using an aggregated Asian group, while controlling for first-generation college student status, 185% poverty threshold, highest math course taken, and taking advanced math coursework.

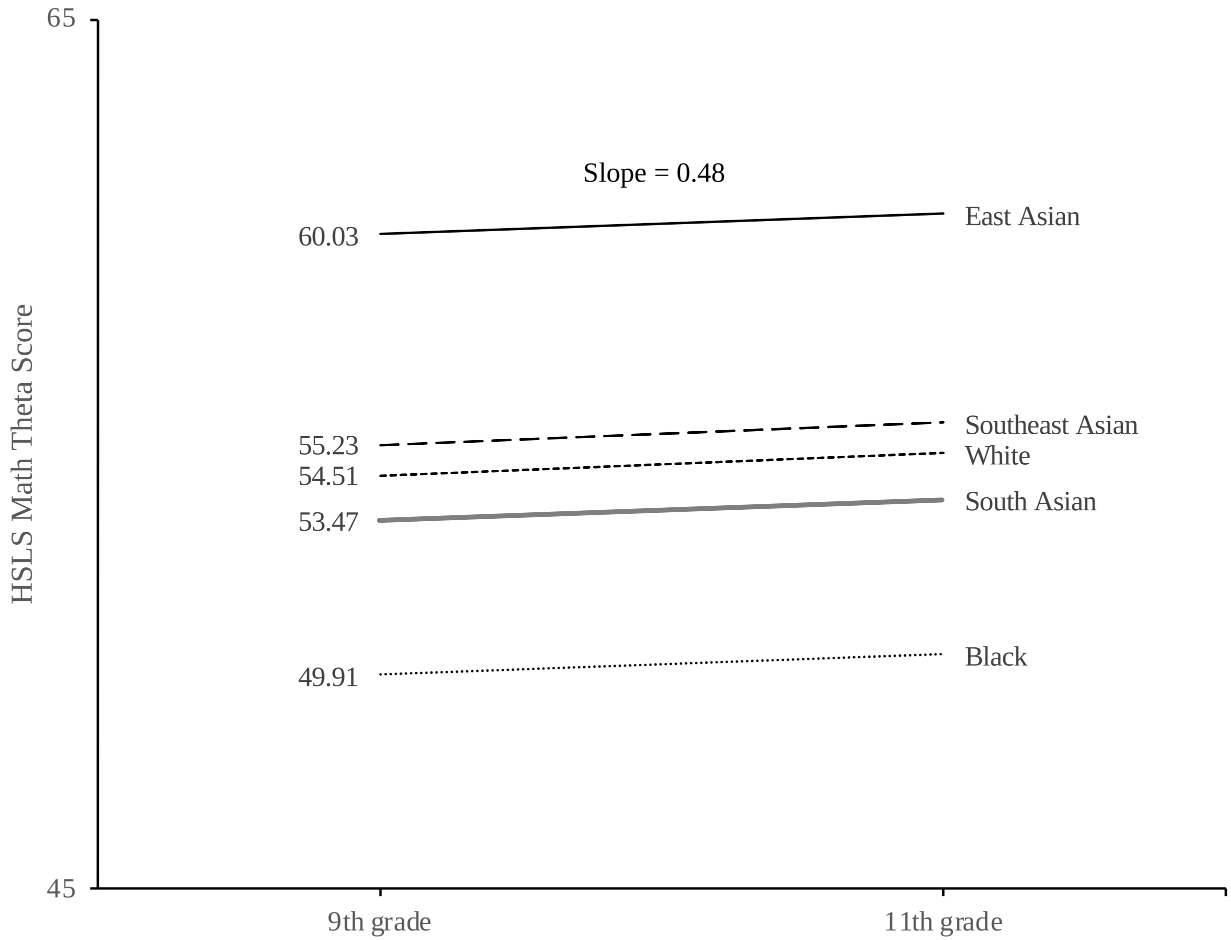


Figure 15. Fitted regression line of the change in HSLS math theta scores from 9th grade to 11th grade by race using disaggregated Asian subgroups, while controlling for first-generation college student status, 185% poverty threshold, highest math course taken, and taking advanced math coursework.

MCA. Model equations are presented in Equations 3 to 10.

Use of an aggregate Asian group

Model A.1: White students as the reference group

$$\begin{aligned} &Math\ Achievement \\ &\quad = 53.68 - 0.45(Grade) + 1.81(Asian) - 6.43(Black) \\ &\quad - 1.35(Female) - 6.38(EnglishLearner) \\ &\quad - 11.70(EligibleFRPL) - 4.49(ReceiveSPED) \end{aligned} \tag{7}$$

Model A.2: Black students as the reference group

$$\begin{aligned} &Math\ Achievement \\ &\quad = 47.25 - 0.45(Grade) + 8.24(Asian) \\ &\quad + 6.43(White) - 1.35(Female) \\ &\quad - 6.38(EnglishLearner) - 11.70(EligibleFRPL) \\ &\quad - 4.49(ReceiveSPED) \end{aligned} \tag{8}$$

Use of disaggregated Asian subgroups

Model B.1: White students as the reference group

$$\begin{aligned} &Math\ Achievement \\ &\quad = 53.66 - 0.45(Grade) + 1.94(AsianEnglish) \\ &\quad + 9.18(ChineseMandarin) + 6.30(OtherEastAsian) \\ &\quad + 4.36(SouthAsian) + 4.30(Vietnamese) \\ &\quad + 1.23(Hmong) + 0.28(OtherLang) - 0.89(Khmer) \\ &\quad - 3.55(Lao) - 5.07(OtherSEA) - 6.54(Black) \\ &\quad - 1.35(Female) - 5.81(EnglishLearner) \\ &\quad - 11.71(EligibleFRPL) - 4.42(ReceiveSPED) \end{aligned} \tag{9}$$

Model B.2: Black students as the reference group

$$\begin{aligned} &Math\ Achievement \\ &\quad = 47.13 - 0.45(Grade) + 1.94(AsianEnglish) \\ &\quad + 15.72(ChineseMandarin) \\ &\quad + 12.84(OtherEastAsian) + 10.90(SouthAsian) \\ &\quad + 10.84(Vietnamese) + 7.77(Hmong) \\ &\quad + 6.82(OtherLang) + 5.65(Khmer) + 2.99(Lao) \\ &\quad + 1.47(OtherSEA) + 6.54(White) - 1.35(Female) \\ &\quad - 5.81(EnglishLearner) - 11.71(EligibleFRPL) \\ &\quad - 4.42(ReceiveSPED) \end{aligned} \tag{10}$$

Race and all covariates were significant predictors of math achievement when using an aggregate Asian group (Model A) and disaggregated Asian subgroups (Model B), including sex, English learner status, receiving special education services, and FRPL status (Table 21). Parameter estimates of fixed effects, variance of random effects, number of estimated parameters, and model fit statistics are provided in Table 22, while fitted line graphs comparing the change in math achievement between 8th grade to 11th grade across race while accounting for covariates are displayed in Figure 16 and 17. On average, math achievement decreases by 0.45 points in both models between 8th and 11th grade. In both models, the fixed effect for the change is statistically significantly different from zero based on the t-test. Further, random intercepts accounted for 77% of the variance of the random effects.

As an aggregate group, Asian students in Model A had significantly higher math achievement by 1.81 points compared to White students (Model A.1) and 8.241 compared to Black students (Model A.2) while accounting for covariates. When compared to White students, results in Model B.1 using disaggregated Asian subgroups show Chinese-Mandarin (β = 9.18), Other East Asian language (β = 6.30), South Asian language (β = 4.36), Vietnamese (β = 4.30), English (β =1.94), and Hmong (β = 1.23) language groups typically performed higher in math achievement while holding demographic covariates constant. Students with Other Southeast Asian language (β = -5.1) and Lao home language (β = -3.55) were the only groups to perform significantly lower in math achievement compared to White students. All Asian language subgroups had significantly higher mean math scores compared to Black students (Model B.2). Students with a Chinese-Mandarin home language had the largest mean difference from Black students of 15.77 points, while students with an Other Southeast Asian home language had the lowest mean difference by 1.47 points. Math achievement decreased on average if students were female (β = -1.35 in both models), received English learner services (β = -6.38 in Model A and β = -5.81 in Model B), received special education services (β = -11.70 in Model A and β = -11.71 in Model B), and were eligible for FRPL (β = -4.49 in Model A and β = -4.42 in Model B).

Table 21
Type III Tests of fixed effects predicting MCA standardized math scores.

	Numerator df	Denominator df	F	*p*-value
Model A: Aggregated Asian group				
Intercept	1	63,463.797	133388.797	0.000
Grade	1	49,443.399	288.074	0.000
Race.AggregateAsian	2	57,703.170	1,372.611	0.000
Female	1	56,468.144	371.532	0.000
EnglishLearner	1	61,249.492	830.139	0.000
ElgibileFRPL	1	58,450.001	10,495.011	0.000
NoSPED	1	56,724.807	2,739.389	0.000
Model B: Disaggregated Asian subgroup				
Intercept	1	60,017.022	40,328.175	0.000
Grade	1	49,469.329	287.025	0.000
Race.DisaggregateAsian	11	58,353.152	279.787	0.000
Female	1	56,491.175	370.702	0.000
EnglishLearner	1	61,591.199	574.688	0.000
ElgibileFRPL	1	58,480.365	10,559.050	0.000
NoSPED	1	56,742.074	2,620.229	0.000

Table 22

Comparison of parameter estimates, number of parameters estimated, error variance, and model fit indices predicting MCA standardized math scores between 8th grade and 11th grade, with Models A.1 and A.2 using an aggregated Asian group and Models B.1 and B.2 using disaggregated Asian groups.

	Aggregated Asian Group		Disaggregated Asian Groups	
	Model A.1 (Race reference group = White)	Model A.2 (Race reference group = Black)	Model B.1 (Race reference group = White)	Model B.2 (Race reference group = Black)
Fixed Effects				
Intercept	53.679***	47.249***	53.660***	47.125***
Slope	-0.450***	-0.450***	-0.449***	-0.449***
Asian	1.811***	8.241***		
Chinese-Mandarin			9.180***	15.717***
South Asian language			4.362***	10.899***
Other East Asian language			6.299***	12.836***
English			1.935***	8.472***
Vietnamese			4.297***	10.834***
Other language			0.280	6.818***
Khmer			-0.890	5.647***
Lao			-3.547***	2.991***
Hmong			1.230***	7.768***
Other Southeast Asian language			-5.071***	1.467*
Black	-6.430***		-6.537***	
White		6.430***		6.537***
Female	-1.351***	-1.351***	-1.347***	-1.347***
English Learner	-6.379***	-6.379***	-5.811***	-5.811***
Eligible for FRPL	-4.492***	-4.492***	-4.418***	-4.418***
Receiving SPED	-11.704***	-11.704***	-11.709***	-11.709***

*$p < .050$, **$p < .010$, ***$p < .000$

Table 22

Comparison of parameter estimates, number of parameters estimated, error variance, and model fit indices predicting MCA standardized math scores between 8th grade and 11th grade, with Models A.1 and A.2 using an aggregated Asian group and Models B.1 and B.2 using disaggregated Asian groups (continued…).

	Aggregated Asian Group		Disaggregated Asian Groups	
	Model A.1 (Race reference group = White)	Model A.2 (Race reference group = Black)	Model B.1 (Race reference group = White)	Model B.2 (Race reference group = Black)
Variance of Random Effects				
Intercepts	58.846	58.846	58.482	58.482
Error	17.182	17.182	17.178	17.178
Number of estimated parameters	10	10	19	19
Model fit statistics				
AIC	708,580.637	708,580.637	708,281.539	708,281.539
BIC	708,676.251	708,676.251	708,463.205	708,463.205

* $p < .050$, ** $p < .010$, *** $p < .000$

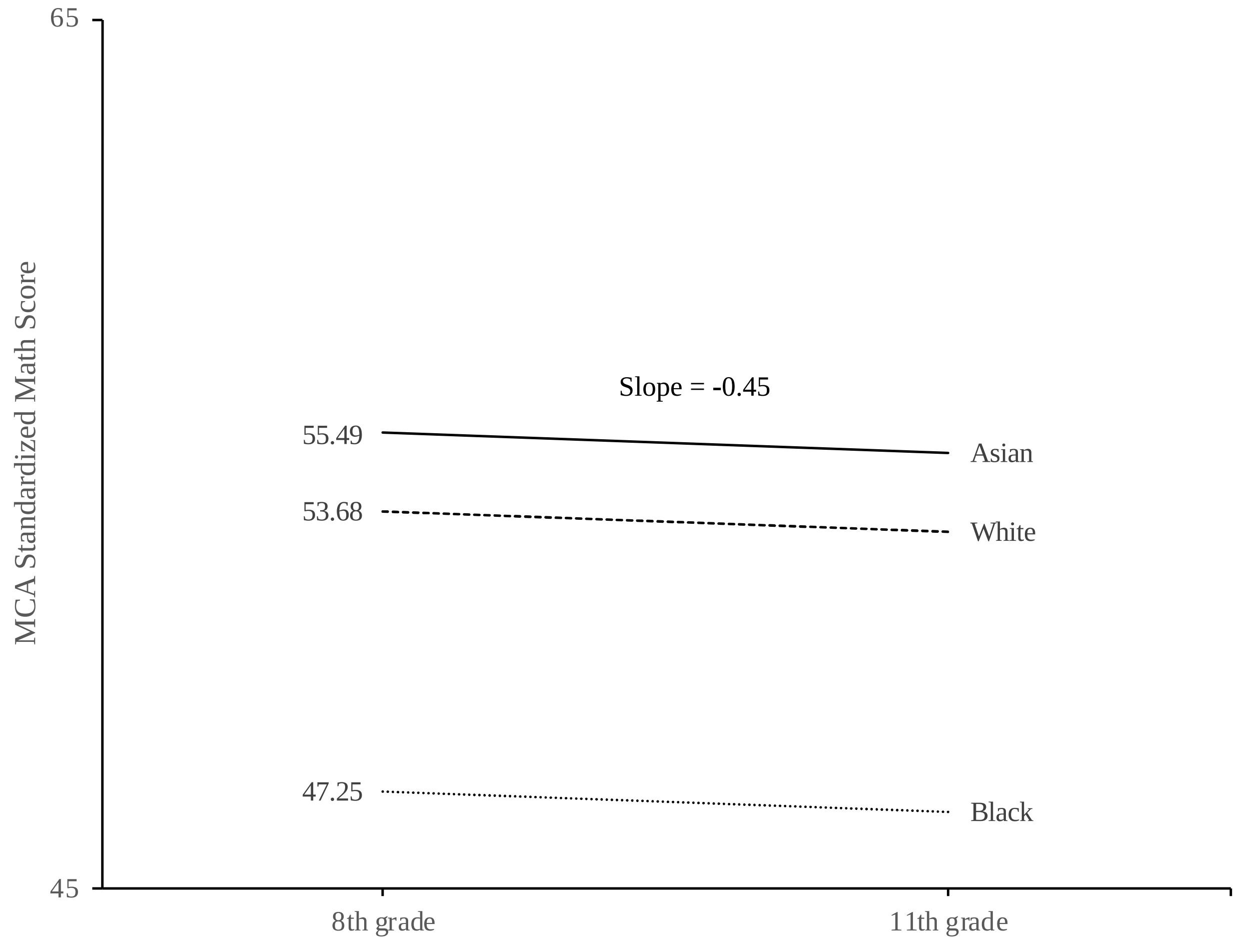


Figure 16. Fitted regression line of the change in MCA standardized math scores from 8th grade to 11th grade by race using an aggregated Asian group, while controlling for sex, English learner status, FRPL status, and receipt of special education services.

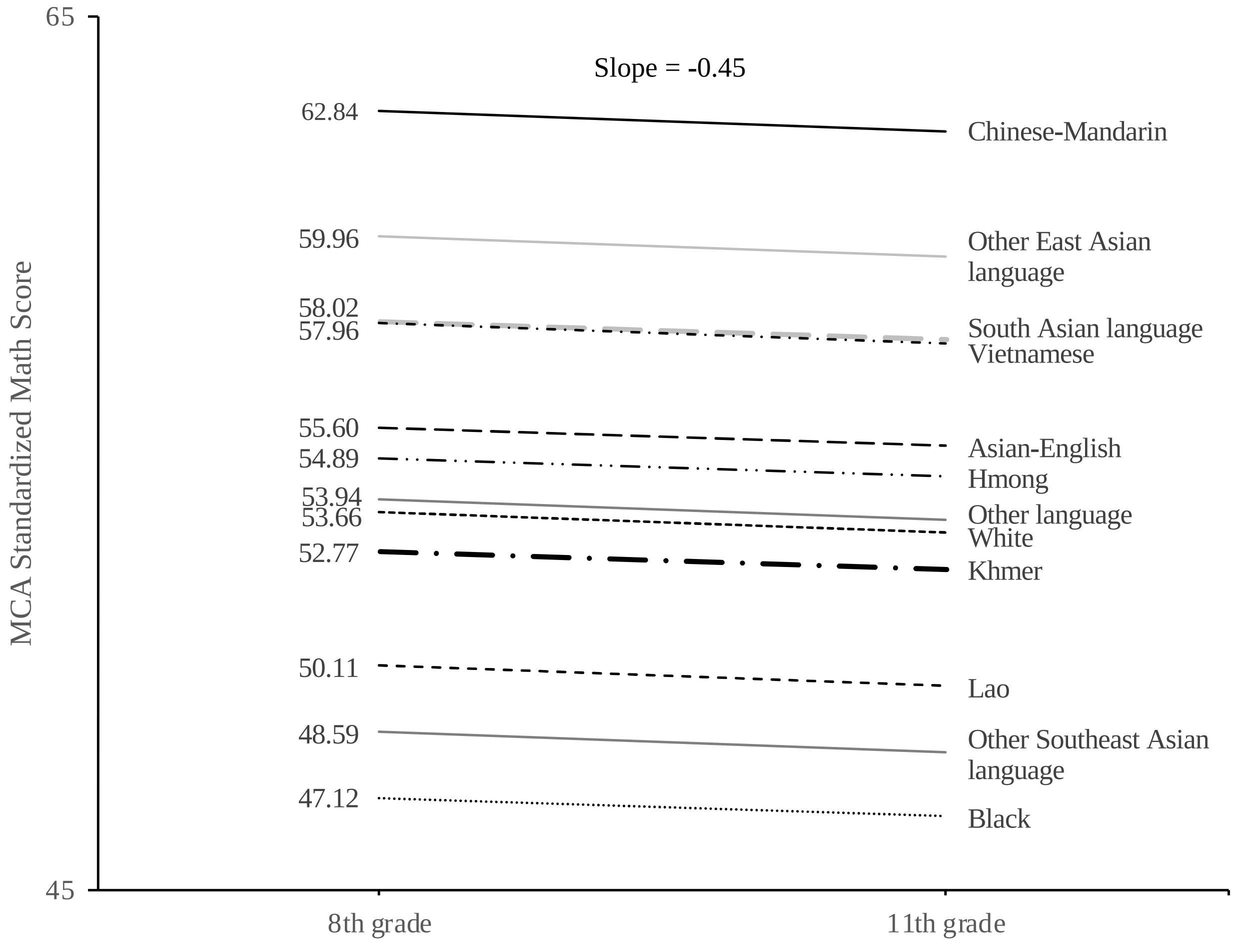


Figure 17. Fitted regression line of the change in MCA standardized math scores from 8th grade to 11th grade by race using disaggregated Asian subgroups, while controlling for sex, English learner status, FRPL status, and receipt of special education services

Chapter 5: Discussion and Implications

The Model Minority Myth (MMM) has depicted AA as homogeneously hard-working, highly educated, and economically successful. The stereotype continues to be pervasive even after more than 50 years since its conception. Educational data have contributed to the credence and spread of the MMM by consistently finding AA students academically outperforming all other race and ethnicity groups. However, researchers and policy advocates have countered the MMM with disaggregated data showing hidden academic disparities across Asian subgroups. This study used two data sets: (1) a nationally representative data set, HSLS, and (2) a state-level data set, MCA-II, to assess math disparities within and across race, while examining differences in using an aggregate Asian group compared to disaggregated Asian subgroups. Analyses focused on the cohort of Asian students, Black students, and White students who were 9th graders attending a public school in the 2009-10 school year.

Discussion of Findings

This study examined three research questions:

1) To what extent are the distributions of math scores for Asian, White, and Black student groups bimodal/multimodal? If there is evidence of multimodality, how do the mode groups differ across demographics?
2) To what extent are there differences in mean math scores across AA subgroups?
3) To what extent are there differences in math scores during high school by race while grouping Asians as an aggregated group compared to disaggregating Asians into subgroups?

Bimodality of math achievement. There was no evidence to suggest that bimodality or multimodality of math achievement is specific to only Asian students. The Silverman Test of Multimodality was used to determine the number of modes present in the distribution of math scores. In the HSLS data, 9th grade math scores were unimodal for Asian and White students, while it was multimodal for Black students. At 11th grade, math scores were bimodal for Asian students and multimodal for White students and Black students. With the MCA data, Silverman Test results indicated the presence of multimodality in 8th and 11th grade for all three race groups. The size of the modes varied relative to the primary mode (i.e., the most frequent mode). The secondary and/or tertiary modes in the HSLS data were moderately sized relative to the primary mode. In the MCA data, the secondary and/or tertiary modes were relatively very small compared to the primary mode—with the exception of modes across 11th grade scores for Black students.

When mode groups were compared within each race group, there were specific demographics that distinguished low and high achievers regardless of race. Low achieving mode groups consistently had more students from low-income households in both the HSLS and MCA data. Further, HSLS data showed low achieving mode groups also had more students who would be the first in their family to attend college. Unfortunately, the MCA data set did not have first-generation college student status available.

There were some demographics that varied more across mode groups for Asian students specifically. First, mode groups among Asian students differed by racial and ethnic subgroup. The low scoring mode groups typically included more Southeast Asian students in both the HSLS and MCA data, while high scoring mode groups consisted of more East Asian students. Detailed subgroup information was not available for Black students and White students in the HSLS dataset, but primary home language was available for all students in the MCA data set. Among Black students, there was some evidence that subgroup is a factor in math disparities within race, where the low-scoring mode group had a higher proportion of Somali students than higher scoring mode groups. However, the vast majority of Black students had an English home language. Secondly, immigrant generation varied more among Asian students than White students and Black students. In the HSLS data, the vast majority of mode groups among Asian students were second-generation immigrants, whereas nearly all students across mode groups were third-generation immigrants or later among Black students and White students. Third, English proficiency also varied more across Asian mode groups. Mode groups among Asian students were fairly evenly split across those having an English first language, a non-English first language, or both an English and non-English home language; whereas, nearly all mode groups among Black students and White students had English as their first language. In the MCA data, differences in whether students received English learner services varied more widely across mode groups among Asian students, while the vast majority of mode groups among Black students and almost all students across mode groups among White students were not English learners.

There is very limited research on the bimodality of achievement scores among AA students. Teranishi (2010) has descriptively examined the distribution of SAT scores between Asians and Whites, while Hartlep, Morgan, and Hodge (2015) examined underlying subpopulations among Asian students in the ELS:02 national data set. Findings from this study are consistent with those studies in that there was a presence of bimodality or multimodality in scores among Asian students. Low- and high-achieving groups varied across subgroup, family poverty level, first-generation college student status, and language proficiency differed. Thus, Asian students are not a homogeneous group. But this research also provides evidence that neither White students nor Black students are comprised of a homogeneous group. What is unknown, however, is whether low- and high-achieving groups differ by racial and ethnic subgroup. Regardless, poverty and first-generation college student status within any race tend to be disproportionate among low-achieving students.

Math disparities across Asian subgroups. ANOVA results revealed evidence of math achievement disparities across Asian subgroups in both the HSLS and MCA data supporting the hypothesis of Southeast Asian students typically underperforming other Asian subgroups. In the HSLS data set, mean differences in math achievement were not statistically different in 9th or 11th grade across the original five Asian subgroups (Chinese, Filipino, South Asian, Southeast Asian, and Other Asian) using the recommended $\alpha = 0.010$ level accounting for sampling design effects, but were significant at the $\alpha = 0.050$ level. Chinese students had statistically significant higher math scores than Southeast Asian students and Filipino students. The five Asian subgroups were recombined into three groups to be mutually exclusive based on geographic region since Filipinos are Southeast Asian. After this reclassification, results show Southeast Asian students had significantly lower mean math scores in 11th grade than East Asian students at the $\alpha = 0.010$ level. Differences in 9th grade mean math scores were significant at the $\alpha = 0.050$ level. South Asian students had a mean math score in-between that of East Asians and Southeast Asians, where mean differences were not statistically significant from either group.

In the MCA data, ANOVA and post-hoc test results also supported the hypothesis of Southeast Asian students having lower levels of math achievement compared to other Asian subgroups. Four of the five Southeast Asian language subgroups, including Khmer, Hmong, Lao, or Other Southeast Asian language groups, had significantly lower mean math scores than their peers with a Chinese-Mandarin, South Asian, Other East Asian language, or English home language. Students with a Vietnamese home language had a significantly lower mean math score than those with a Chinese-Mandarin or South Asian home language, but had a significantly higher mean math score compared to their Southeast Asian peers.

Disparities across Asian subgroups in this study were found consistent with prior research (Council on Asian Pacific Minnesotans, 2012; Her, 2014; Lee et al., 2017; Lee, 1994; Lee, 1996; Lee & Zhou, 2015; Ngo & Lee, 2007; Pang et al., 2011; Portes & Rumbaut, 2001; Tamura, 2003; Teranishi, 2010; Tran & Birman, 2010; Yang, 2004; Yang, 2013). When examining mean scores across Asian subgroups without accounting for any background characteristics, results generally show that Chinese students typically have the highest math achievement levels whereas South Asian students closely follow and Southeast Asian students have the lowest levels of achievement. There are important considerations surrounding the low achievement of Southeast Asian students, however. First, Southeast Asian study samples in previously mentioned studies have largely included war refugees or children of war refugees predominantly from Cambodia, Laos, and Vietnam. These students typically come from families with low economic backgrounds and have parents with little or no experience with formal education. And yet, Southeast Asians can be comprised of those from many other countries, including Brunei, East Timor, Indonesia, Thailand, the Philippines, Singapore, Malaysia, Myanmar, and other small islands in the geographical region. Differences in history, economic development, and political conflicts of Southeast Asian countries have shaped immigration patterns to the U.S. and influence the economic and social conditions of Southeast Asian immigrants in the U.S. Findings from research studies and program evaluations can vary greatly depending on the make-up of Southeast Asian samples or participants.

Secondly, there are disparities even across specific Southeast Asian subgroups. Vietnamese students typically outperform Cambodian, Hmong, and Lao students (Pang et al., 2011; Her, 2014; Portes & Rumbaut, 2001; Rumbaut, 2005). Researchers have noted that Vietnamese refugees immigrating to the U.S. shortly after the end of the Vietnam War were highly educated with some English proficiency and professional skills (Kula & Paik, 2017; Lee, 2015; Ngo & Lee, 2007; Portes & Rumbaut, 2001). Their resettlement in the U.S. was concentrated in California, where they were able to establish ethnic enclaves that helped the adjustment and resettlement of other Vietnamese refugees with limited education, English, and work skills coming after them. On the other hand, many Lao and Hmong refugees lacked formal education and were resettled across the U.S. in cities without established Asian communities. Zhou and Bankston's (1994) study focusing on Vietnamese students found that having strong ties to and involvement in their ethnic community is a form of social capital that positively impacts the adaptation and academic success of second-generation Vietnamese students. Upon closer examination, having more Vietnamese friends than White or Black friends was a stronger predictor of academic achievement among Vietnamese students than family structure or parental educational background (Bankston, Caldas, & Zhou, 1997).

Third, Southeast Asians can also include ethnic Chinese students with parents from Vietnam and Cambodia depending on how they choose to identify themselves. Chinese students can be a very diverse group, including those from families who have been in the U.S. for many generations, who have immigrant parents from China or another Asian country (e.g. Taiwan, Singapore, South Korea), and those with refugee parents from Vietnam and Cambodia. Few studies have mentioned or examined this complexity (Kim, 2002; Lee, 1994), but it can contribute to the variation in academic achievement among Southeast Asian students and Chinese students. Kim's (2002) study controlled for Chinese ethnicity when examining academic differences between Vietnamese and Cambodian students, who were 1.5 or 2nd generation immigrants. Vietnamese students performed higher than Cambodian students by 5.8 percentile points in reading and 14.2 percentile points in math after controlling for student demographics, cultural variables, and family structural variables. However, if students were ethnically Chinese, their reading score increased by 10.8 percentile points on average and math scores increased by an average of 8.8 percentile points, implying that Chinese students from Vietnam and Cambodia have higher levels of achievement than their non-Chinese counterparts on average. Yet in Lee's (1994) qualitative study of high- and low-achieving Asian students, three interviewees who identified as ethnically Chinese from Cambodia or Vietnam struggled with reading and writing, were not engaged in school, and/or lacked motivation to do well in school. The economic and educational backgrounds within the

Southeast Asian and Chinese communities can vary greatly and have an influence on students' academic achievement.

Examination of aggregated vs. disaggregated groups for Asian students. The hypothesis of Southeast Asian students having significantly lower levels of math achievement compared to White students, but having similar levels of math achievement as Black students was not fully supported. In both the HSLS and MCA analysis, LME modeling showed that using an aggregated Asian group did conceal disparities across disaggregated subgroups even after controlling for selected characteristics. There were also different patterns in the disparities when subgroups were compared to White students and Black students.

In the HSLS analysis, using an aggregated Asian race group was consistent with the widely known finding of Asian students outperforming White students and Black students. Eight covariates were included in the initial model, but only four were statistically significant at the $\alpha = 0.010$ level, including 185% poverty threshold, first-generation college student status, highest math course taken, and advanced math course taking. Asian students had a higher mean score than White students after controlling for the four covariates, but the mean difference was only statistically significant at the $\alpha = 0.050$ level and not at the recommended $\alpha = 0.010$ level to account for sampling design effects. Asian students scored significantly higher than Black students on average given all four covariates at the $\alpha = 0.010$ level. When Asian students were disaggregated using three groups, results showed that East Asian students had a significantly higher mean math score between 9th and 11th grade than White students and Black students at the $\alpha = 0.010$ level. Southeast Asian students also had a higher mean math score than White students and Black students between grades, but the mean score difference was only statistically significant from that of Black students at the $\alpha = 0.010$ level. As is, this finding contradicts that Southeast Asian students have lower achievement levels than White students. However, data on parental birthplace for these Southeast Asian students reveal that most (63%) had at least one parent born in Vietnam. Several studies have found that Vietnamese students typically outperform White students in math (Bankston, Caldas, & Zhou, 1997; Council on Asian Pacific Minnesotans, 2012; Pang et al., 2011). This highlights how the make-up of AA students, even Southeast Asian students, can contribute to inconsistent findings and inferences of AA student achievement. South Asian students had a slightly lower mean math score than White students and a higher mean math score compared to Black students, but neither mean difference in scores were statistically significant. This finding is inconsistent with findings from the MCA analysis and prior studies as well. South Asian students typically outperform White students and most other Asian subgroups as well, with the exception of Chinese students in some instances (Kao, 1995; Pang et al., 2011; Yang, 2013). This is finding is surprising because most students (65%) in this group had at least one parent born in India and Asian Indians tend to outperform White students (Kao, 1995; Pang et al., 2011). However, the estimates for South Asian students may be unreliable given the very small effective sample size of 40 students compared to 70 Southeast Asian students and 2,330 White students.

Dual language and immigrant generation were not significant predictors of math achievement given the other covariates. Dual language, i.e., the first language students learned to speak, was used as a proxy for English proficiency but is a weak proxy to use. Further, it could have been insignificant because the HSLS study excludes students who were deemed as not having enough English proficiency to complete the math assessment (Ingel et al., 2011). In prior studies, Asian students with foreign-born parents (i.e., first- and second-generation immigrant children) tended to have higher academic achievement than those with U.S.-born parents (i.e., third or later generations) (Duong, Badaly, Liu, Schwartsz, & McCarty, 2016: Galindo & Pong, 2014; Kao & Rutherford, 2007; Zhang, 2001). The advantage lies in high parental expectations and aspirations for their children, as well as a strong ethnic identity (Harris et al., 2008; Portes & Rumbaut, 2001; Rumbaut, 2005; Sue & Okazaki, 1990). However, previous research studies had not examined immigrant generation in combination with math coursework data. Student background characteristics may be of little use in explaining further variation in achievement above and beyond coursework. Poverty level and first-generation college student status, two covariates that consistently contribute to achievement gaps, were not statistically significant at the $\alpha=0.050$ level when all eight covariates were included in the LME modeling. These two covariates were not significant predictors of math achievement until sex and immigrant generation were removed from the model.

Findings suggest that the type and the rigor of math coursework are highly correlated with math achievement and are important to assess racial achievement gaps in mathematics. Students who took higher level math courses in 11th grade, like trigonometry, pre-calculus, calculus, or statistics, had higher math achievement than students who took algebra II/III, which was the most common level of math taken among students. Students who took advanced math coursework also had higher math achievement than those who did not. Asian students, specifically East Asian students, had the highest mean math theta score compared to their peers, but there is not enough information in the LME modeling to suggest they are more likely to take higher levels and/or more rigorous math coursework. Yet results from the examination of modes in the distribution of math theta scores in 11th grade (i.e., research question 1) shed some light in further investigating racial achievement gaps and math coursework. Among Asian students in the high scoring mode group (i.e., a score of 60 or greater), the vast majority of students have taken either pre-calculus (42%) or calculus (37%) and half of the students have taken advanced math coursework (51%). Furthermore, nearly half of these students were East Asian, identifying their subgroup as Chinese (32%) or other Asian (12%), such as Korean or Japanese. In comparison, the majority of White students and Black students in a similar scoring mode group (e.g., a score of greater than 60) had taken algebra II/III (19% and 38% respectively) or pre-calculus (49% and 44% respectively). Almost one-third of White students (32%) had taken advanced math coursework, while only 22% of Black students had done so. Thus, there is some evidence to suggest that Asian students take higher level and more rigorous math coursework compared to White students and Black students. And, there appears to be some differences in math coursework across Asian subgroups with East Asian students taking higher level math courses and more advanced math coursework than South Asian students and Southeast Asian students.

These findings are consistent with prior studies suggesting the sequence and rigor of math courses help to explain math achievement gaps across race (Byun, Irvin, & Bell, 2015; Davenport et al., 2013). Davenport et al. (2013) found that the amount and type of math courses have significant impacts on math achievement throughout high school while controlling for 8th grade math achievement. Asian students, on average, had the highest math achievement level, took the highest number of courses, and progressed through the “optimal” sequence of math courses compared to students of other racial groups. Black students, who had the second lowest mean math achievement level just above American Indian students, took more math courses than expected, but did not progress through higher levels of math. Results showed that the type of math course was more predictive of math achievement than the number of courses taken. Yet it was unclear why course-taking patterns differed across race. Davenport et al. (2013) suggest students should be recommended to take the highest level math given their ability.

In the MCA analysis, results with an aggregate Asian race group showed that Asian students outperformed both White students and Black students after controlling for all available covariates: sex, English learner status, eligibility for FRPL, and receiving SPED services. With disaggregated Asian subgroups, there were only two Asian subgroups with significantly lower mean math scores compared to White students, students with a Lao home language by 3.5 points and students with an Other Southeast Asian home language by 5.1 points. Mean score differences between the other eight Asian subgroups compared to White students ranged between 0.3 for students with an Other home language to 9.2 points for students with a Chinese-Mandarin home language. All Asian subgroups had significantly higher mean math scores than Black students, with a wide range of mean score differences with the Other Southeast Asian group having a higher mean score of about 1.5 points and Chinese-Mandarin students having a mean score of about 15.7 points higher.

Generally, the results of this study affirm arguments made by researchers and advocates that aggregated data are misleading in describing the achievement of AA students. There were variations in the Asian-White and Asian-Black achievement gaps after using disaggregated subgroups. In the HSLS analysis, the achievement gap widened between East Asians with Whites and Blacks, while the gaps for Southeast Asians and South Asians compared to White and Blacks narrowed. A similar pattern occurred in the MCA analysis as well. Achievement gaps with Whites and Blacks widened for half of the language groups (Chinese-Mandarin, Other East Asian, South Asian, Vietnamese, and English), while gaps narrowed for the other half (Hmong, Khmer, Other language, Lao, and Other Southeast Asian). The use of disaggregated groups slightly improved the fit of the model compared to the use of an aggregated group according to model fit statistics.

The demand for using disaggregated subgroup data for Asian students is rooted in creating equitable opportunities for those subgroups who do not fit the MMM, in particular Southeast Asians. However, most Asian subgroups were not at a disadvantage compared to White students or Black students after controlling for selected covariates. Past researchers have specifically found Cambodian, Lao, and Hmong students trailing behind White students (Her, 2014, Pang et al., 2011; Teranishi, 2010). The MCA analysis suggests that Hmong students in MN are faring very similar to Whites in math achievement and do much better than Black students on average. Lao and Other Southeast Asian students had significantly lower levels of math achievement on average compared to White students and the other Asian subgroups, suggesting these two subgroups are experiencing difficulties. Unfortunately, there is not enough information to determine why Lao students have lower levels of math achievement. In preliminary analyses, Lao students did not differ in demographics compared to students from the other Asian subgroups. Students with an Other Southeast Asian language consisted of primarily of Burmese students (69%). Burmese students in MN during this time were recent refugee arrivals from Myanmar. Overall, the Southeast Asian language group was predominantly English learners (77%), which likely explains their lower levels of math achievement. English learners typically have lower achievement levels than their English proficient peers (Lee, 1994; Tandon, 2016). Further, English proficiency is a source of anxiety for students, often leading to feelings of self-doubt, shamefulness, and social isolation. In Tandon's (2016) qualitative study, Burmese refugee students expressed feelings of shame to ask for help, but at the same time, perceived English proficiency as necessary to pursue college and enter the workforce.

Implications of Findings

This study and its findings contribute to academic and policy debates around the use of disaggregated subgroup data. While this study focused on the AA population, implications can apply to all race groups. The premise behind calls for using disaggregated subgroup data is to dispel the MMM by more accurately assessing achievement gaps for AA subgroups, but also to advocate for Southeast Asian subgroups that do not fit the MMM. Aggregated data on AA students have consistently suggested there is no need for support by showing that an achievement gap compared to Whites does not exist. When decisions are made from these results, some Southeast Asian subgroups are left invisible and overlooked for support and opportunities. For example, AA are not considered an underrepresented minority group in higher education and in the STEM workforce, where many of the opportunities are filled by international students and H1-B visa workers from China and India. AA are often not targeted or recruited for programs and opportunities to increase minority representation at colleges or STEM fields. And yet, disaggregated data show that some Southeast Asian subgroups are underrepresented and experience similar social and economic barriers that formally recognized underrepresented groups do. Activists argue it is problematic to use an aggregated Asian group in accessing academic achievement because it is not reflective of all subgroups. It is not a matter of some individuals doing better than others, but disparities pertain to cultural communities that go ignored. This study corroborated that argument where the mean math achievement for AA students as an aggregate group was not reflective of the mean achievement for the different subgroups. In particular, the MCA analysis showed statistically significant achievement gaps for Lao and Burmese students, who need more academic support to meet the achievement levels of their other Asian and White peers. More specifically, Burmese students need more support in gaining English proficiency. While Burmese and Lao students need additional support, their need is not as great as it is with Black students. Results suggest Black students are experiencing difficulties in making comparable achievement to Asian students and White students. And while White students are typically used as the control group to assess disparities, it does not mean that certain subpopulations do not experience challenges and struggle academically.

Efforts to close racial achievement gaps are reinforced by civil rights laws with the implication that minority students would have fair opportunities for higher education and work when they have equivalent levels of achievement as White students. In other words, racial minority students' knowledge and skills will be competitive without regard to race. However, that is under the assumption that institutional racism and discrimination do not exist. Racial minority students experience discrimination that is hard to capture in measuring and assessing achievement gaps. For AA students specifically, a misconception of the MMM is that AA do not experience racism and discrimination (Museus & Kiang, 2009). As prevalent as the MMM is, AA are also commonly seen as foreigners (Kawai, 2005; Ng et al. 2007; Wang, 2008; Yu, 2006). This perception of AA also has historical ties to xenophobic media campaigns in the late 1800s leading up to the Chinese Exclusion Act of 1882, before the conception of the MMM, where AA were dehumanized and projected as the "yellow peril" overtaking the U.S. and threatening the availability of jobs. This stereotype has manifested in more present times as well. A common experience for AA is being asked, "Where are you from?" with the expectation for an Asian country rather than a U.S. city or state. Additionally, the derogatory racial epithet aimed at AA, as well as people of color and immigrants, is to "go back to where you came from," asserting that the U.S. is not a home for AA.

There is a movement in efforts to remedy racial achievement gaps through equity rather than equality under critical race theory (Howard & Navarro, 2016; Singleton, 2014; Tate, 1997). This means that supports are targeted and provided in culturally relevant ways that meet the needs of specific racial groups with the acknowledgment that race impacts how students are treated within systems. There is potential in the use of disaggregated race and ethnic subgroup data to examine disparities that are seen in experiences, but not reflected in data and research. This may be what is fueling the demands for disaggregated data.

The need for disaggregated data grows also because the data currently available are flawed for use. Disaggregated race and ethnicity subgroup data are hard to come by. Census data contain the most detailed disaggregated race and ethnic subgroup data, however, its use for research on achievement gaps among students is limited because data are by household. Very few national longitudinal studies collect race and ethnic subgroup data. As mentioned previously and shown in Table 2, four studies do so. Still, it is often only available in restricted-use files limited access for its use. Further, the subgroup categories available are inconsistent across studies. “Asian Indian” is used in two earlier studies, while later studies use “South Asian.” “Vietnamese” is also a category available in the earlier two studies, while they are grouped with “Southeast Asian” in others in the later ones. And “Korean” and “Japanese” are specifically used in three studies, while they are grouped with “Other Asian” in the remaining study. The inconsistency in what subgroup categories are available and used lead to convoluted research findings that make it difficult to understand disparities and make decisions.

In an effort to be more consistent, as previously mentioned, a handful of local state governments are starting to or exploring the collection and reporting of disaggregated race and ethnicity data. State agencies and education departments are using U.S. Census detailed race and ethnicity as guides for which subgroup options should be made available based on their population. Yet data disaggregation bills have been difficult to pass. These legislation bills also face heavy criticism, and in states where such bills have been passed, there is risk of it being eliminated every year.

The way AA are classified is also of concern. With limited access to subgroup data, common proxies like primary home language or birthplace are used. These proxies are becoming less and less reliable. Take for instance the “English” home language among AA students in the MCA data. This group can include adopted children from Asia (e.g., Korea or China), students who have lost or have declining proficiency in their native language, and/or those whose parents do not want them to be singled out for English learner services for being bilingual. Minnesota has the largest concentration of Korean adoptees in the world (Choy, 2016; Jackson, Kim, Nelson, & Huie, 2010; Nelson, 2017). Further, there is evidence of declining use of and fluency in the Hmong language among second-generation and third-generation Hmong youth (Bosher, 1997; Vang, 2008; Yang, 2008; Withers, 2004). These families may be more likely to choose English as their primary home language than Hmong.

Subgroup category options and labels have also been problematic. In the HSLS data, there were not mutually exclusive groups with both a Filipino subgroup and a Southeast Asian subgroup. While minimal, this is a potential emotional risk for research participants. Filipino students could have felt as being seen as separate from their Southeast Asian peers, which can also have historical connotations, as the Philippines was once a U.S. colony. Conversely, students who chose the Southeast Asian category could have very well asked themselves why their subgroup was not represented and feel ignored. These feelings could challenge a student to reflect on their ethnic identity and how they are perceived by researchers, as well as the mainstream society.

As the use of disaggregated race and ethnicity data grows, it is important for researchers, evaluators, and policy makers to reflect on what decisions will be made from inferences on individuals and to consider unintended consequences (Kane, 2001; Messick, 1995). Inferences made about which AA subgroups are “low achievers” and which subgroups are “high achievers” will impact decisions made about which groups receive supports. Using Kane’s (2001) criteria for examining consequences, the inferences made about Asian individuals from this study include 1) Chinese students are the high achievers and 2) Lao and Burmese students are low achievers. Resulting decisions from these inferences mean some AA subgroups will benefit, while other subgroups will not. In particular, the Chinese community has been very vocal about the pitfalls of using disaggregated data, primarily around discriminatory policies and unequal treatment that has revived historical trauma from decades of racist policies (e.g., the Chinese Exclusion Act of 1882, segregated Chinese schools, and the Immigration Act of 1924 forbidding immigrants from Asia) (Fuchs, 2016; Fuchs, 2017a; Fuchs, 2017b; Moser, 2018; Wang, 2016b; Wang, 2017). Chinese students will be seen as fitting the MMM and perceived to be high achievers who do not need help or face discrimination based on their high levels of academic success. This could lead to Chinese students being overlooked or having to fight harder for opportunities. A predominant concern among Chinese protesters of data disaggregation is college admissions to Ivy League schools and affirmative action policies. Chinese parents are concerned their children will have to meet higher expectations than other races to be admitted, and that their children will face increased competition in being overlooked in preference for Southeast Asian students and other racial minority students who typically have faced more economic hardships. These types of unintended consequences are very real and emotionally charged. It is difficult to navigate, but researchers, evaluators, and policy makers need to be prepared for the discussion.

Based on the results of this study, the following are important implications for researchers and evaluators:

1) Collect disaggregated racial and ethnic subgroup data when it will be useful in interpreting the findings and the decision-making process. There

is a demand for disaggregated subgroup data by social justice advocates and researchers alike. Data have been disaggregated in many different ways—not just by race and ethnicity—to examine achievement gaps and find ways to remedy them. But first, disaggregated data must be available for use. Equally important, researchers and evaluators need to use consistent subgroup labels. The U.S. Census has the most comprehensive list of detailed race and ethnic groups (Panapasa, Crabbe, & Kaholokula, 2011), which should be used as a reference for determining which subgroups data should be collected for based on population and which subgroup labels to use. As the collection and use of detailed disaggregated subgroup data become more consistent, there will no longer be a need to use proxies like home language, national origin, and birthplace to identify subgroup. These proxies are increasingly becoming less and less accurate.

2) Allow individuals to self-select detailed race and ethnicity subgroup with an "Other" write-in option. Personal demographics are almost always asked for in research and evaluation. It is important to ensure that study subjects and program participants see themselves reflected in the categories provided and are given an opportunity to choose for themselves which group they identify with. An "Other" option can produce varied results and be unreliable and unactionable, however, the priority is for subjects and participants to be treated with respect and to ensure there are no or minimal risks to their involvement in research and evaluation.

3) Use mutually exclusive race categories and/or subgroup categories with "check all that apply" option. In the HSLS student survey, the categories Filipino and Southeast Asian were offered, but the two groups are not exclusive. This potentially puts individuals in an unpleasant situation where they may feel ignored, misrepresented, or misunderstood. In addition, race questions are typically asked with a "check all that apply" option, but it does not appear that Asian subgroup question was asked in this way. The HSLS codebook and questionnaires do not specifically reference this question as a "check all that apply" format like it does with the question asking about race. Further, the HSLS data set only provides one composite variable for Asian subgroup. Thus, it is assumed that students could only choose one option.

4) Respect Asian and Native Hawaiian/Pacific Islanders (NHPI) as separate racial groups. NHPI students are often grouped with Asian students because of small sample sizes. Researchers and evaluators should respect these two groups as distinct and refrain from combining these groups together. When faced with this dilemma, researchers and evaluators should consider the advantages and disadvantages of either excluding the group with the small sample size altogether or at the very least acknowledging the combination and noting the research of disparities between these groups.

5) Know the subgroup make-up of study samples, use subgroup labels accordingly, and provide detailed information. When study samples are

comprised of entirely of a specific subgroup, use the detailed subgroup label rather than the all-encompassing "Asian" label. For example, use "Cambodian" rather than "Asian." Additionally, when Asian study samples include a mix of subgroups, report the frequencies of those subgroups for context given the known disparities.

6) Invest in accommodating language needs of Asian study subjects. English language proficiency is a key determinant of academic success and proficiency varies within the Asian community. Researchers and evaluators need to invest in appropriately accommodating language needs in both written and verbal form, such as translating data collection materials, ensuring academic assessments are readable and culturally relevant, and using native speaker interpreters.

7) Examine disparities across racial and ethnic subgroups when possible to determine whether the overall mean is reflective of subgroups. As researchers and evaluators, time is spent on running preliminary analyses to check the quality of data and assumptions for statistical significance testing. When race and ethnic data are available, running an ANOVA to examine subgroups should be a part of the preliminary analyses process especially when diverse groups will be represented by a mean score. If statistically significant disparities exist across race and ethnic subgroups, it important to use the disaggregated subgroups rather than proceed with using the aggregated race and ethnic group.

8) Acknowledge disparities across subgroups. There is a large body of research showing disparities across Asian subgroups and the damaging effects of the MMM. When an aggregate Asian group is the only option, it is important that economic and academic disparities are at the very least acknowledged to prevent the prevalence of the MMM.

9) Be aware of community perceptions about disaggregating data by racial and ethnic subgroup. There is growing advocacy for using disaggregated data for Asian Americans, but there is not a consensus for it within the Asian community. The Asian community is divided on the issue with concerns about how disaggregated data will be used to advantage or disadvantage specific subgroups in accessing limited resources and opportunities. It is important for researchers and evaluators to be aware and do their due diligence to keep up with current events and news surrounding the debate. This will help in understanding the context of how findings will be received and used by community stakeholders. Further, perceptions about disaggregating subgroups are likely to vary across racial groups given differences in historical events and political sentiments.

This study also has implications for policy makers and program decision makers:

1) Acknowledge the misleading MMM and recognize differences within the Asian community. The MMM continues to be pervasive and can cloud decision-making processes that leave AA largely ignored. This study showed that AA students are diverse, especially in family poverty levels,

first-generation college student status, immigrant generation status, and English proficiency.

2) Do not use national-level data to make local-level decisions. Results from the national-level HSLS data set were very different from state-level MCA data. HSLS data did not reveal concerning disparities between Asian subgroups and White students, however, MCA data revealed disparities showing that students with a Lao or an Other Southeast Asian home language had lower math achievement than White students.

3) Consult with cultural community members in interpreting data and findings. Advocates and members of a cultural community can help provide historical, social, and political context to data and study findings. For example, this study showed that Asian subgroups differed by family poverty level, first-generation college student status (i.e., parental educational background), and English proficiency. The wide range of these characteristics within the AA community can be attributed to historical immigration patterns and national ethnic origins.

4) Increase opportunities and participation in math interventions and programs for students who need the support. Needs should be assessed on an individual basis. Lao, Burmese, and Black student groups had lower levels of math achievement on average compared to students from other Asian subgroups and White students. These specific groups need intensive math support to make enough growth to close the achievement gaps, such as a second math support course, math intervention programs, and/or at

the very least tutoring and homework help. Additionally, as Asian students commonly display passive behaviors in the classroom, Lao and Burmese students who are struggling academically may need some social-emotional support to learn how to advocate for themselves and seek help.

5) Increase academic opportunities, interventions, and programming for English learners. This study showed that bilingualism and English proficiency varied more so among AA students than White students and Black students. Those who lacked English proficiency were overrepresented in the low-scoring achievement groups in the MCA analysis. English proficiency is needed to “make it” in the U.S., both socially and economically. This issue is also not specific to the AA community; thus, increased efforts to support English learner students would beneficial for all students who need the support. In addition, English learner service programs should be evaluated to assess the quality of services, the process of how students enter and exit the program, and the impact of programming on students’ academic growth.

Limitations

This study is not without limitations. First, the data used for this study are nearly 10 years old. Findings are likely not a reflection of the current state of disparities. Secondly, the effective sample size for Asian subgroups in the HSLS data set were quite small relative to sample sizes for White students and Black students. Small sample sizes impact the reliability of estimates, especially in comparing groups. Third, the analysis of MCA data did not have as many covariates to use as the HSLS data set. Math course taking data were not available in the MCA data set for the vast majority of students and MDE does not collect first generation college student status or immigrant generation data. Fourth, this study did not standardize MCA scores to assess proficiency levels as it is typically used by MDE to assess the extent to which students are proficient on state standards. Thus, while there are differences in scores across subgroups and race, it does not tell whether one group is more or less proficient on average compared to another. Fifth, the potential to observe increases in MCA standardized scores across time are limited. In particular, students with high scores are limited from making as much gain as those with low scores. Sixth, SPSS was used to complete LME modeling as it was within the author's capabilities and it appropriately handled sampling weights. In doing so, the scaled identity covariance structure was the only option for having two assessment points. This covariance structure assumes constant variance and no correlation between errors of measures within an individual, but errors between multiple measures for an individual tend to be highly correlated. Lastly, the analyses only tested the main effects of each covariate. In the HSLS analysis, the main effects of 185% poverty threshold and first-generation immigrant status were not statistically significant predictors of math achievement at the $\alpha = 0.050$ given all else in the model until the covariates, sex and immigrant generation, were removed. The LME modeling did not test the interactions between these terms; thus, it is unknown whether the relationship between these variables are influential in assessing disparities in math achievement.

Future Research

There are many opportunities for future research to examine the use of disaggregated race and ethnic subgroups. There is a need for more research and evaluation to use disaggregated data for Asian students. The demand for its use grows among social justice and policy advocates. To date, there is not enough research available to completely rule out its collection and use in assessing academic disparities. There is also too limited research on which to draw definite conclusions given the lack of and inconsistent manner in which disaggregated data has been collected. Research and evaluation using disaggregated Asian subgroup data have disproven the MMM, yet its impact on dispelling widespread beliefs in AA as the model minority is incomplete. Moreover, there is an urgent need for qualitative research to better understand the divide within the AA community regarding disaggregated subgroup data. The concerns of both sides of the debate are valid and must be better understood to meet needs, address concerns, and alleviate unintended consequences within the Asian community. Future studies should also focus on specific subgroups, like the Lao and Burmese, to identify specific barriers and types of supports needed. This study focused on access to and use of disaggregated data to examine the presence of disparities across Asian subgroups, but more in-depth research is needed to understand what would be helpful to close disparities. Further, there should be efforts to track the progress of these two groups over time to assess the impact of supports on academic achievement.

Additional research is needed to assess whether disaggregated data by subgroup is relevant in other race and ethnicity groups. In the MCA data, there was some evidence that Somali students, who are primarily refugee children or children of refugees, typically have lower math achievement among black students. Yet researchers have typically found that foreign-born African American immigrants and their children (including immigrants from Africa and the Caribbean) tend to have higher academic success than Black students born in the U.S. (Giraldo-Garcia & Bagaka, 2017; Hudley, 2016; Pinder, 2012; Tauriac & Liem, 2012). In addition, subgroup data is rarely collected for White students. It is unknown whether subgroup is an important factor in academic achievement disparities within the White community, as well as the extent to which socioeconomic status and parental education attainment is disproportionate across subgroups.

More research is needed to examine disparities in math course-taking patterns. In the HSLS analysis, the type of math course taken and whether students took any advanced math courses (i.e., AP, IB, or math courses granting college credits), were strong predictors of math achievement. There was some evidence to suggest that Asian students, particularly East Asian students, take higher levels of math courses like calculus and were more likely to advanced math coursework compared to White students, Black students, and other Asian subgroups. Further study is needed to assess how students are advised and choose their math coursework in high school, as well as how race plays into that process. Future studies on AA students and math coursework should also examine the extent of differences across racial and ethnic subgroups as well.

Future studies should examine the extent of subgroup disparities in reading and writing. This study focused on math achievement where MCA results showed English proficiency was a significant predictor of math achievement. In both the HSLS and MCA datasets, English proficiency or dual language was also an important factor in distinguishing low and high achievers among AA students. Given the wide variation of English language proficiency within the Asian community, research on subgroup disparities in reading can help to identify needed supports and evaluate the extent to which available programming and interventions are working to close achievement gaps across groups.

Further research is also needed to examine patterns of achievement across more than two time points. This study did not have enough assessment periods to perform LME with random slopes. In other words, there was not enough information with two time points to examine individual changes across time in achievement. Future research should assess whether there are disparities across race and subgroups in patterns of academic achievement across time. The MCA data showed there was an average decline in students' math achievement from 8^{th} to 11^{th} grade, but specific groups could very well have experienced an increase while other groups may have had more dramatic decrease in achievement.

Conclusion

This study contributes to the body of research around academic disparities within the AA community and provides guidance on the use of disaggregated race and ethnicity data. Certainly, no racial group is ever completely homogenous. However, it is hard to ignore the prevalence of the MMM clouding perceptions of AA students and negatively impacting their psychosocial development and opportunities for equitable support. Advocacy groups are passionately demanding disaggregated data to track academic progress for vulnerable and/or special interest subgroups. Researchers and evaluators have, to some extent, the control to decide what data to collect and how to collect it based on inquiry demands. They also have the responsibility to ensure that the use of data, findings, and inferences are valid and that there is a conscious effort to consider potential unintended consequences.

References

ACT. (2018). *The ACT profile report – National: Graduating class of 2017.* Iowa City: ACT. Retrieved from http://www.act.org/content/act/en/research.html

Adereth, M. (2014, October 12). Silverman's mode estimation method explained [Blog post]. Retrieved from http://adereth.github.io/blog/2014/10/12/silvermans-mode-detection-method-explained/

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

American Immigration Council. (2018, April 6). The H-1B visa program: A primer on the program and its impact on jobs, wages, and the economy. Retrieved from https://www.americanimmigrationcouncil.org/research/h1b-visa-program-fact-sheet

Bankston III, C. L., Caldas, S. J., & Zhou, M. (1997). The academic achievement of Vietnamese American students: Ethnicity as social capital. *Sociological Focus*, *30*(1), 1-16.

Barrett, A. N., Barile, J. P., Malm, E. K., & Weaver, S. R. (2012). English proficiency and peer interethnic relations as predictors of math achievement among Latino and Asian immigrant students. *Journal of Adolescence*, *35*(6), 1619-1628.

Beede, D., Julian, T., Khan, B., Lehrman, R., McKittrick, G., Langdon, D., & Doms, M. (2011). *Education supports racial and ethnic equality in STEM. Economics and Statistics Administration*. Washington DC: U.S. Department of Commerce:

Retrieved from http://www.esa.doc.gov/Reports/education-supports-racial-and-ethnic-equality-stem

Bosher, S. (1997). Language and cultural identity: A study of Hmong students at the postsecondary level. *Tesol Quarterly*, *31*(3), 593-603.

Byars-Winston, A., Estrada, Y., Howard, C., Davis, D., & Zalapa, J. (2010). Influence of social cognitive and ethnic variables on academic goals of underrepresented students in science and engineering: a multiple-groups analysis. *Journal of counseling psychology*, *57*(2), 205.

Byun, S. Y., Irvin, M. J., & Bell, B. A. (2015). Advanced math course taking: Effects on math achievement and college enrollment. *The Journal of Experimental Education*, *83*(4), 439-468.

Center for Immigration Studies. (1995, September 1). *The legacy of the 1965 Immigration Act: Three decades of mass immigration.* Retrieved from https://cis.org/Report/Legacy-1965-Immigration-Act

Chang, B., & Au, W. (2007). You're Asian, how could you fail math?: Unmasking the myth of the model minority. In N. D. Hartlep (Ed), *The model minority stereotype reader: Critical and challenging readings for the 21st century* (pp. 11-20). San Diego, CA: Cognella Academic Publishing.

Change the Equation. (2015). *The diversity dilemma: Changing the face of the STEM workforce*. Washington, DC: Change the Equation. Retrieved from http://changetheequation.org/solving-diversity-dilemma

Choy, C. C. (2016). A History of Asian International Adoption in the United States. In D. K. Yoo and E. Azuma (Eds), *The Oxford Handbook of Asian American History*, pp. 205-221. New York, NY: Oxford University Press.

Chu, S. P. L. (2002). *Internalization of the model minority stereotype and its relationship to psychological adjustment* (Doctoral dissertation, ProQuest Information & Learning).

Council on Asian Pacific Minnesotans. (2012). *Asian Pacific students in Minnesota: Facts, not fiction.* Saint Paul, MN. Retrieved from https://mn.gov/capm/resources/council-reports/

Davenport, E. C., Davison, M. L., Wu, Y., Kim, S., Kuang, H., Kwak, N., Chan, C., Ayodele, A. (2013). Number of courses, content of coursework, and prior achievement as related to ethnic achievement gaps in mathematics. *Journal of Educational Leadership in Action, 2*(1). Retrieved from http://www.lindenwood.edu/academics/beyond-the-classroom/publications/journal-of-educational-leadership-in-action/

Department for Professional Employees. (2015). *The professional and technical workforce: Fact sheet 2015*. Washington DC: Department for Professional Employees. Retrieved from http://dpeaflcio.org/programs-publications/issue-fact-sheets/the-professional-and-technical-workforce/

Dillman, D. A., Smyth, J. D. & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method*. Hoboken, NJ: John Wiley & Sons, Inc.

Dinh, Quyen. (2013). *Moving beyond the "Asian" check box*. Washington, DC: Southeast Asian Resource Action Center (SEARAC). Retrieved from http://www.searac.org/content/education-policy-resource-hub

Doshi, F. (2006, April 27). UC accepts unprecedented number of Asian students. *Daily Bruin*. Retrieved from http://dailybruin.com/2006/04/27/uc-accepts-unprecedented-numbe/

Duong, M. T., Badaly, D., Liu, F. F., Schwartz, D., & McCarty, C. A. (2016). Generational differences in academic achievement among immigrant youths: A meta-analytic review. *Review of Educational Research*, *86*(1), 3-41.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2012). *Applied longitudinal analysis*. Hoboken, NJ: John Wiley & Sons.

Fry, R. & Passel, J. S. (2014). *In post-recession era, young adults drive continuing rise in multi-generational living*. Washington, DC: Pew Research Center. Retrieved from http://www.pewsocialtrends.org/2014/07/17/in-post-recession-era-young-adults-drive-continuing-rise-in-multi-generational-living/

Fuchs, C. (2016, August 26). California data disaggregation bill sparks debate in Asian-American community. *NBC News*. Retrieved from https://www.nbcnews.com/news/asian-america/california-data-disaggregation-bill-sparks-debate-asian-american-community-n638286

Fuchs, C. (2017a, June 22). Bill to disaggregate Asian American, Pacific Islander data passes New York Assembly. *NBC News.* Retrieved from https://www.nbcnews.com/news/asian-america/bill-disaggregate-asian-american-pacific-islander-data-passes-new-york-n775556

Fuchs, C. (2017b, July 27). Rhode Island data disaggregation law revives debate among Asian Americans. *NBC News.* Retrieved from https://www.nbcnews.com/news/asian-america/rhode-island-data-disaggregation-law-revives-debate-among-asian-americans-n786986

Fuchs, C. (2018, March 15). Connecticut bill would bar state from collecting student data from specific sub groups. *NBC News.* Retrieved from https://www.nbcnews.com/news/asian-america/connecticut-bill-would-bar-state-collecting-student-data-specific-sub-n857121

Galindo, C., & Pong, S. L. (2011). Tenth grade math achievement of Asian students: Are Asian students still the "Model Minority"?—A comparison of two educational cohorts. In N. D. Hartlep (Ed.), *The model minority stereotype reader: Critical and challenging readings for the 21st century* (pp. 21-48). San Diego, CA: Cognella Academic Publishing.

Giraldo-Garcia, R. J., & Bagaka's, J. G. (2017). Critical Analysis of the Educational Success of African Immigrants and African Americans in the US. *Journal of Global Initiatives: Policy, Pedagogy, Perspective*, *11*(2). Article 2.

Guey, L. & Lubin, G. (2013, Jun 1). Asian Americans are smarter, richer, and harder-working than everyone else. *Business Insider,* Retrieved from http://www.businessinsider.com/the-rise-of-asian-americans-charts-2013-6

Hall, C. C. I. (2009). Asian American women: The nail that sticks out is hammered down. In N. Tewari and A. N. Alvarez (Eds.), *Asian American psychology: Current Perspectives* (pp. 193-209), New York: NY: Psychology Press/Taylor & Francis Group.

Hall, P., & York, M. (2001). On the calibration of Silverman's test for multimodality. *Statistica Sinica*, 515-536.

Hahs-Vaughn, D. L. (2005). A primer for using and understanding weights with national datasets. *The Journal of Experimental Education, 73*(3), 221-248.

Hahs-Vaughn, D. L. (2006). Analysis of data from complex samples. *International Journal of Research and Method in Education, 29*(2), 165-183.

Harris, A. L., Jamison, K. M., & Trujillo, M. H. (2008). Disparities in the educational success of immigrants: An assessment of the immigrant effect for Asians and Latinos. *The ANNALS of the American Academy of Political and Social Science*, *620*(1), 90-114.

Hartlep, N. D. (2014). Introduction – History of the model minority stereotype in the United States. In N. D. Hartlep (Ed.), *The model minority stereotype reader: Critical and challenging readings for the 21st century* (pp. XXI-XXVIII). San Diego, CA: Cognella Academic Publishing.

Hartlep, N. D., Morgan, G. B., & Hodge, K. J. (2015). An Asian American subgroup analysis of the restricted-use ELS 2002 dataset. In N. D. Hartlep and B. J. Porfilio (Eds), *Killing the Model Minority Stereotype: Asian American counterstories and complicity*, (pp. 357-380). Charlotte, NC: Information Age Publishing, Inc.

Her, C. S. (2014). Ready or not: The academic college readiness of Southeast Asian Americans. *Multicultural Perspectives*, *16*(1), 35-42.

Hing, J. (2012, June 21). Asian Americans Respond to Pew: We're not your model minority. *Colorlines,* Retrieved from http://www.colorlines.com/articles/asian-americans-respond-pew-were-not-your-model-minority

Hixson, L., Hepler, B. B., & Kim, M. O. (2012). *The Native Hawaiian and other Pacific Islander population: 2010*. Suitland, MD: U.S. Census Bureau. Retrieved from https://www.census.gov/2010census/data/2010-census-briefs.php

Hoeffel, E. M., Rastogi, S., Kim, M. O., and Shahid, H. (2012). *The Asian Population: 2010*. Suitland, MD: U.S. Census Bureau. Retrieved from https://www.census.gov/2010census/data/2010-census-briefs.php

Holland, A. T., & Palaniappan, L. P. (2012). Problems with the collection and interpretation of Asian-American health data: omission, aggregation, and extrapolation. *Annals of epidemiology*, *22*(6), 397-405.

Howard, T. C., & Navarro, O. (2016). Critical race theory 20 years later: Where do we go from here?. *Urban Education*, *51*(3), 253-273.

Hudley, C. (2015). Achievement and Expectations of Immigrant, Second Generation, and Non-immigrant Black Students in US Higher Education. *Adolescent Identity and Schooling: Diverse Perspectives*, 52

Humes, K., & Hogan, H. (2009). Measurement of race and ethnicity in a changing, multicultural America. *Race and Social Problems*, *1*(3), 111.

Hune, S., & Chan, K. (1997). Special focus: Asian Pacific American demographic and educational trends. In D. Carter & R. Wilson (Eds.), *Minorities in higher education: Fifteenth annual status report: 1996-1997* (pp. 39-59). Washington, DC: American Council on Education.

Islam, N. S., Khan, S., Kwon, S., Jang, D., Ro, M., & Trinh-Shevrin, C. (2010). Methodological issues in the collection, analysis, and reporting of granular data in

Asian American populations: historical challenges and potential solutions. *Journal of health care for the poor and underserved*, *21*(4), 1354.

Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Denver, J. A., Ottem, R., … & Leinwand, S. (2011). *High School Longitudinal Study of 2009 (HSLS:09): Base-year data file documentation (NCES 2011-328).* Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2011328

Ingels, S. J., Pratt, D. J., Herget, D. R., Denver, J. A., Burns Fritch, L., Ottem, R., … & Chistopher, E. (2013). *High School Longitudinal Study of 2009 (HSLS:09): First follow-up data file documentation (NCES 2014-361).* Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2014361

Institute of International Education. (2018a). "International students totals by place of origin, 2012/13-2017/18." *Open Doors Report on International Educational Exchange.* Retrieved from http://www.lie.org/opendoors

Institute of International Education. (2018b). "International students by field of study, 2015/2016 – 2017/18." *Open Doors Report on International Educational Exchange.* Retrieved from http://www.lie.org/opendoors

Jackson, K., Lee, H., Kim, J. R., Nelson, K. P., & Huie, W. Y. (2010). *Here: A Visual History of Adopted Koreans in Minnesota*. Yeong & Yeong Book Company.

Kalton, G. (2009). Methods for oversampling rare subpopulations in social surveys. *Survey methodology*, *35*(2), 125-141.

Kane, M. T. (2001). Current concerns in validity theory. *Journal of educational Measurement*, *38*(4), 319-342.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1-73.

Kao, G. (1995). Asian Americans as model minorities? A look at their academic performance. *American Journal of Education, 103*(2), 121-159.

Kao, G., & Rutherford, L. T. (2007). Does social capital still matter? Immigrant minority disadvantage in school-specific social capital and its effects on academic achievement. *Sociological Perspectives*, *50*(1), 27-52.

Kara, J. (2018, March 14). A debate over how government should identify our ethnicity. *The CT Mirror*. Retrieved from https://ctmirror.org/2018/03/14/debate-government-identify-ethnicity/

Kawai, Y. (2005). Stereotyping Asian Americans: The dialectic of the model minority and the yellow peril. *The Howard Journal of Communications, 16*(2), 109-130.

Kim, R. Y. (2002). Ethnic differences in academic achievement between Vietnamese and Cambodian children: Cultural and structural explanations. *Sociological Quarterly*, *43*(2), 213-235.

Kim, P. Y., & Lee, D. (2014). Internalized model minority myth, Asian values, and help-seeking attitudes among Asian American students. *Cultural Diversity and Ethnic Minority Psychology, 20*(1), 98-106.

Kula, S. M., & Paik, S. J. (2016). A historical analysis of Southeast Asian refugee communities: Post-war acculturation and education in the US. *Journal of Southeast Asian American Education and Advancement*, *11*(1), 1-23.

Lai, E., & Arguelles, D. (2003). *The New Face of Asian Pacific America*. Los Angeles, CA: Asian Week.

Lee, C. (2015). Family reunification and the limits of immigration reform: Impact and legacy of the 1965 Immigration Act. *Sociological Forum, 30*(1), 528-548.

Lee, D. M., Duesbery, L., Han, P. P., Thupten, T., Her, C. S., & Pang, V. O. (2017). Academic needs and family factors in the education of Southeast Asian American students: Dismantling the model minority myth. *Journal of Southeast Asian American Education and Advancement*, *12*(2).

Lee, S. J. (1994). Behind the Model-Minority stereotype: Voices of high- and low-achieving Asian American Students. *Anthropology & Education Quarterly, 25*(4): 413-429.

Lee, S. J. (1996). *Unraveling the "Model Minority" stereotype: Listening to Asian American Youth*. New York, NY: Teachers College Press.

Lee, J., & Zhou, M. (2015). *The Asian American Achievement Paradox*. New York, NY: Russell Sage Foundation.

Li, C. (2018, April 27). Minnesota SF 2597 bill and disaggregated data collection. *China Insight*. Retrieved from https://www.chinainsight.info/opinion-mainmenu/1238-minnesota-sf-2597-bill-and-disaggregated-data-collection.html

Li, G. (2005). Other people's success: Impact of the" model minority" myth on underachieving Asian students in North America. *KEDI Journal of Educational Policy*, *2*(1).

Linn, R. L. (1997). Evaluating the validity of assessments: The consequences of use. *Educational Measurement: Issues and Practice*, *16*(2), 14-16.

Marpsat, M., & Razafindratsima, N. (2010). Survey methods for hard-to-reach populations: Introduction to the special issue. *Methodological Innovations Online*, *5*(2), 3-16.

Martin, M. (2012, June 20). Asian-Americans on the rise. *National Public Radio,* Retrieved from http://www.npr.org/2012/06/20/155431944/asian-americans-on-the-rise

McCarthy, T. (2012, June 19). Asian immigrants to America surpass Hispanics in 2011 numbers. *The Guardian*. Retrieved from https://www.theguardian.com/world/2012/jun/19/asian-immigrants-america-surpass-hispanics

Mehrens, W. A. (1997). The consequences of consequential validity. *Educational Measurement: Issues and Practice*, *16*(2), 16-18.

Messick, S. (1989). Validity. In Linn, R. L. (Ed.), *Educational measurement, (3rd ed.)* (pp. 13-103). New York, NY: American Council on Education, Macmillan Publishing Company.

Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational measurement: Issues and practice*, *14*(4), 5-8.

Moser, E. (2018, March 9). Asian groups disagree on student classifications. *The Day*. Retrieved from https://www.theday.com/article/20180309/NWS01/180309318

Munroe, T. (n.d.). The great American success story – the rise of Asian Americans. *The Catholic Business Journal,* Retrieved from http://catholicbusinessjournal.biz/content/great-american-success-story%E2%80%94-rise-asian-americans

Museus, S. D. (2009). A critical analysis of the exclusion of Asian American from higher education research and discourse." In L. Zhan (ed.), *Asian American Voices: Engaging, Empowering, Enabling* (pp. 59–76). New York: NLN Press.

Museus, S. D. (2014). *Asian American students in higher education.* New York, NY: Routledge.

Museus, S. D., & Chang, M. J. (2009). Rising to the challenge of conducting research on Asian Americans in higher education. *New Directions for Institutional Research*, *2009*(142), 95-105.

Museus, S. D., & Kiang, P. N. (2009). Deconstructing the model minority myth and how it contributes to the invisible minority reality in higher education research. *New Directions for Institutional Research*, *2009*(142), 5-15.

Museus, S. D., & Liverman, D. (2010). High-performing institutions and their implications for studying underrepresented minority students in STEM. *New Directions for Institutional Research*, *2010*(148), 17-27.

Museus, S. D., Maramba, D. C., & Teranishi, R. T. (2013). Introduction. In S. D. Museus, D. C. Maramba, and R. T. Teranishi (Eds.), *The misrepresented minority: New insights on Asian Americans and Pacific Islanders, and the implications for higher education* (pp. 3-11). Sterling, VA: Stylus Publishing, LLC.

Museus, S. D., Palmer, R. T., Davis, R. J., & Maramba, D. (Eds.). (2011). Racial and Ethnic Minority Student Success in STEM Education [Special issue]. *ASHE Higher Education, 36*(6).

National Commission on Asian American and Pacific Islander Research in Education (CARE). (2010). *Federal higher education policy priorities and the Asian*

American and Pacific Islander Community. Los Angeles, CA: CARE. Retrieved from http://care.gseis.ucla.edu/care-reports/

National Commission on Asian American and Pacific Islander Research in Education (CARE). (2013). *iCount: A data quality movement for Asian Americans and Pacific Islanders in Higher Education*. Los Angeles, CA: CARE. Retrieved from http://care.gseis.ucla.edu/care-reports/

National Forum on Education Statistics. (2016). *Forum Guide to Collecting and Using Disaggregated Data on Racial/Ethnic Subgroups*. (NFES 2017-017). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/forum/pub_2017017.asp

National Science Foundation (NSF) & National Center for Science and Engineering Statistics (NCSES). (2017). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2017.* Special Report NSF 17-310. Arlington, VA: NSF & NCSES. Available at www.nsf.gov/statistics/wmpd/

Nelson, K. P. (2017). Korean transracial adoption in Minnesota. *MNOPEDIA*. Retrieved from http://www.mnopedia.org/korean-transracial-adoption-minnesota

Ng, J.C., Lee, S. S., & Pak, Y. K. (2007). Contesting the model minority and perpetual foreigner stereotypes: A critical review of literature on Asian Americans in Education. *Review of Research in Education, 31*(1), 95-130.

Ngo, B., & Lee, S. J. (2007). Complicating the image of model minority success: A review of Southeast Asian American education. *Review of educational research*, *77*(4), 415-453.

Nord, C., Roey, S., Perkins, R., Lyons, M., Lemanski, N., Brown, J., & Schuknecht, J. (2011). *The Nation's Report Card: America's High School Graduates* (NCES 2011-462). U.S. Department of Education & National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Office of Management and Budget (OMB). (1997, October 30). Revisions to the standards for the classification of federal data on race and ethnicity: Federal register notice. Washington, DC: Executive Office of the President. Retrieved from https://www.whitehouse.gov/omb/fedreg_1997standards

Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment*, *7*(3), 367.

Osborne, J. W. (2011). Best practices in using large, complex samples: The importance of using appropriate weights and design effect compensation. *Practical Assessment, Research and Evaluation, 16*(2), 1-7.

Pak, Y. K., Maramba, D. C., & Hernandez, X. J. (Eds.). (2014). *Asian Americans in Higher Education: Charting New Realities: AEHE Volume 40, Number 1*. John Wiley & Sons.

Panapasa, S., Crabbe, K. O., & Kaholokula, J. K. A. (2011). Efficacy of federal data: revised Office of Management and Budget standard for Native Hawaiian and other Pacific Islanders examined. *AAPI Nexus: Policy, Practice and Community*, *9*(1-2), 212-220.

Pang, V. O., Han, P. P., & Pang, J. M. (2011). Asian American and Pacific Islander students: Equity and the achievement gap. *Educational Researcher*, *40*(8), 378-389.

Park, C. C. (2000). Learning style preferences of Southeast Asian students. *Urban Education*, *35*(3), 245-268.

Petersen. W. (1966, January 9). Success story, Japanese American style. *The New York Times*, pp. 20-21, 33, 36, 38, 40-41, 43.

Pew Research Center. (2013). *The Rise of Asian Americans, Updated edition*. Washington, DC. Retrieved from http://www.pewsocialtrends.org/2012/06/19/the-rise-of-asian-americans/

Pew Research Center. (2018). *Income inequity in the U.S. is rising most rapidly among Asians*. Washington, DC. Retrieved from http://www.pewsocialtrends.org/2018/07/12/income-inequality-in-the-u-s-is-rising-most-rapidly-among-asians/

Pinder, P. J. (2012). Afro-Caribbean and African American Students, Family Factors, and the Influence on Science Performance in the United States: The Untold Story. *Education*, *132*(4).

Popham, W. J. (1997). Consequential validity: Right Concern-Wrong Concept. *Educational measurement: Issues and practice*, *16*(2), 9-13.

Portes, A., & Rumbaut, R. G. (2001). *Legacies: The story of the immigrant second generation*. Berkeley, CA: University of California Press.

Pratt, B. M., Hixson, L., Jones, N. A. (2015). *Measuring race and ethnicity across the decades: 1790-2010, Mapped to 1997 U.S. Office of Management and Budget Classification Standards (Infographic)*. Suitland, MD: U.S. Census Bureau. Retrieved from https://www.census.gov/data-tools/demo/race/MREAD_1790_2010.html

Ramakrishnan, K. & Ahmad, F. Z. (2014). *State of Asian Americans and Pacific Islanders Series: A multifaceted portrait of a growing population*. Washington, DC: Center for American Progress. Retrieved from https://www.americanprogress.org/issues/race/reports/2014/04/23/87520/state-of-asian-americans-and-pacific-islanders-

Reckase, M. D. (1998). Consequential validity from the test developer's perspective. *Educational Measurement: Issues and Practice*, *17*(2), 13-16.

Redford, J. & Hoyer, K. M. (2017). *First-generation and continuing-generation college students: A comparison of high school and post-secondary experiences*. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018009

Reeves, R. V. & Halikias, D. (2017, February 1). Race gaps in SAT scores highlight inequality and hinder upward mobility. *Brookings*. Retrieved from https://www.brookings.edu/research/race-gaps-in-sat-scores-highlight-inequality-and-hinder-upward-mobility/

Ruiz, N. G. (2017, April 27). *Key facts about the U.S. H-1B visa program*. Washington, DC: Pew Research Center. Retrieved from http://www.pewresearch.org/fact-tank/2017/04/27/key-facts-about-the-u-s-h-1b-visa-program/

Rumbaut, R. (2004). Ages, life stages, and generational cohorts: Decomposing the immigrant first and second generations in the United States. *International Migration Review, 38*(3), 1160-1205.

Rumbaut, R. G. (2005). Children of immigrants and their achievement: the roles of family, acculturation, social class, gender, ethnicity, and school context.

Shadish, W. R., Cook, T. D., Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Belmont, CA: Wadsworth Cengage Learning.

Shen, F. C., Wang, Y. W., & Swanson, J. L. (2011). Development and initial validation of the internalization of Asian American stereotypes scale. *Cultural Diversity and Ethnic Minority Psychology*, *17*(3), 283.

Shepard, L. A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, *16*(2), 5-24.

Siegel, L. (2012, October 27). Rise of the Tiger Nation. *The Wall Street Journal,* Retrieved from http://www.wsj.com/articles/SB10001424052970204076204578076613986930932

Silverman, B. W. (1981). Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society. Series B (Methodological)*, 97-99.

Singleton, G. E. (2014). *Courageous conversations about race: A field guide for achieving equity in schools*. Corwin Press.

"Success Story of one minority group in the U.S." (1966, December 26). *U.S. News & World Report*. pp. 73-76.

Sue, S., & Okazaki, S. (1990). Asian-American educational achievements: A phenomenon in search of an explanation. *American psychologist*, *45*(8), 913.

Suzuki, B. H. (1977). Education and the socialization of Asian Americans: A revisionist analysis of the "Model Minority" thesis. *Amerasia, 4*(2), 23-51.

Snyder, T. D. (2018). *Mobile Digest of Education Statistics, 2017* (NCES 2018-138). U.S. Department of Education, Washington, DC. National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018138

Tauriac, J. J., & Liem, J. H. (2012). Exploring the divergent academic outcomes of US-origin and immigrant-origin Black undergraduates. *Journal of diversity in higher education*, *5*(4), 244.

Tamura, E. H. (2003). Introduction: Asian Americans and educational history. *History of Education Quarterly*, *43*(1), 1-9.

Tandon, M. (2016). Resettlement Struggles of Burmese Refugee Students in US High Schools: A Qualitative Study. *Journal of Southeast Asian American Education & Advancement*, *11*(1).

Tang, M. (2008). Psychological impacts of "model minority" on Asian Americans. *Model minority myth revisited: An interdisciplinary approach to demystifying Asian American educational experiences*, 117-132.

Tate IV, W. F. (1997). Critical race theory and education: History, theory, and implications. *Review of research in education*, *22*(1), 195-247.

Teranishi, R. T. (2010). *Asians in the ivory tower: Dilemmas of racial inequity in American higher education. Multicultural Education Series.* New York, NY: Teachers College Press.

Thomas, S. L. and Heck, R. H. (2001). Analysis of large-scale secondary data in higher education research: Potential perils associated with complex sampling designs. *Research in Higher Education, 42*(5), 517-540.

Thompson, T. L., & Kiang, L. (2010). The model minority stereotype: Adolescent experiences and links with adjustment. *Asian American Journal of Psychology*, *1*(2), 119.

Thorndike, R. L., & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education (8th ed.)*. New York City, NY: Pearson Education, Inc.

Tran, N., & Birman, D. (2010). Questioning the model minority: Studies of Asian American academic performance. *Asian American Journal of Psychology*, *1*(2), 106.

U.S. Census Bureau. American Community Survey, 2017 American Community Survey 1-Year Estimates, Table B02001, retrieved from American FactFinder: http://factfinder.census.gov

U.S. Census Bureau. American Community Survey, 2017 American Community Survey 1-Year Estimates, Table B02015, retrieved from American FactFinder: http://factfinder.census.gov

U.S. Census Bureau. American Community Survey, 2017 American Community Survey 1-Year Estimates, Table S0201, retrieved from American FactFinder: http://factfinder.census.gov

U.S. Citizenship and Immigration Services. (n.d.). *Preference system (Immigration Act of 1990)*. Retrieved from https://www.uscis.gov/tools/glossary/preference-system-immigration-act-1990

U.S. Department of Education. (2016, November 14). *U.S. Department of Education awards $836,000 to 3 states for Asian American Pacific Islander Data Disaggregation Initiative* [Press release]. Retrieved from https://www.ed.gov/news/press-releases/us-department-education-awards-836000-3-states-asian-american-pacific-islander-data-disaggregation-initiative

Vang, N. (2008, May 5). Hmong language endangered?. *Twin Cities Daily Planet*. Retrieved from https://www.tcdailyplanet.net/hmong-language-endangered/

Wang, F. (2016a, May 9). New data disaggregation initiative announced by Dept. of Education. *NBC News*. Retrieved from http://www.nbcnews.com/news/asian-america/new-data-disaggregation-initiative-announced-dept-education-n568856

Wang, F. (2016b, August 10). California's Asian-Pacific Islander community divide over race data bill. *ABC 10 (Connect)*. Retrieved from http://www.abc10.com/news/politics/divided-asian-pacific-islander-community-over-controversial-bill/293968887

Wang, H. L. (2017, August 5). 'Racist Bill'?: Chinese immigrants protest effort to collect more Asian-American data. *National Public Radio News*. Retrieved from https://www.npr.org/2017/08/05/541844705/protests-against-the-push-to-disaggragate-asian-american-data

Wang, L. L. (2008). Myths and realities of Asian American success: Reassessing and redefining the "Model Minority" stereotype. In G. Li & Wang, L. (Eds.), *Model minority myth revisited: An interdisciplinary approach to demystifying Asian American educational experiences* (pp. 21-42). Charlotte, NC: Information Age Publishing, Inc.

Watanabe, P. Y. (2015) Asian Americans Rise Up: The Response to the Pew Report on *The Rise of Asian Americans*. *AAPI Nexus: Policy, Practice and Community, 13*(1-2), 321-338.

West, B. T., Welch, K. B., & Galecki, A. T. (2015). *Linear mixed models: a practical guide using statistical software, 2nd edition*. Chapman and Hall/CRC: Boca Raton, FL.

Withers, A. C. (2004). Hmong language and cultural maintenance in Merced, California. *Bilingual Research Journal*, *28*(3), 425-461.

Wong, F., & Halgin, R. (2006). The "model minority": Bane or blessing for Asian Americans?. *Journal of Multicultural Counseling and Development*, *34*(1), 38-49.

Wright, W. E., & Boun, S. (2011). Southeast Asian American education 35 years after initial resettlement: Research report and policy recommendations. *Journal of Southeast Asian American Education and Advancement*, *6*(1), 1.

Yang, K. (2004). Southeast Asian American Children: Not the "Model Minority". *Future of Children*, *14*(2), 127-133.

Yang, L. (2013). Educational Achievement Among Asian Children: Ethnic Differences in First Grade Math and Reading Scores. *McNair Scholars Research Journal*, *9*(1), 16.

Yang, T. (2008). Hmong Parents Critical Reflections on Their Children's Heritage Language Maintenance. *Journal of Southeast Asian American Education and Advancement*, *3*(1), 17.

Yoo, H. C., Burrola, K. S., & Steger, M. F. (2010). A preliminary report on a new measure: Internalization of the Model Minority Myth Measure (IM-4) and its psychological correlates among Asian American college students. *Journal of Counseling Psychology*, *57*(1), 114.

Yoo, H. C., Miller, M. J., & Yip, P. (2015). Validation of the internalization of the Model Minority Myth Measure (IM-4) and its link to academic performance and psychological adjustment among Asian American adolescents. *Cultural Diversity and Ethnic Minority Psychology*, *21*(2), 237.

Yu, T. (2006). Challenging the politics of the "model minority" stereotype: A case for educational equality. *Equity & Excellence in Education*, *39*(4), 325-333.

Zhang. Y. (2001, April). *Immigrant generational differences in academic achievement and its growth: The case of Asian American high school students*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Zhou, M., & Bankston III, C. L. (1994). Social capital and the adaptation of the second generation: The case of Vietnamese youth in New Orleans. *International migration review*, *28*(4), 821-845.

Zong, J. & Batalova, J. (2017a, June 7). *Refugees and asylees in the United States*. Washington, DC: Migration Policy Institute. Retrieved from https://www.migrationpolicy.org/article/refugees-and-asylees-united-states

Zong, J. & Batalova. J. (2017b, September 29). *Chinese immigrants in the United States*. Washington, DC: Migration Policy Institute. Retrieved from https://www.migrationpolicy.org/article/chinese-immigrants-united-states