Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1486

Hmong but not Asian, Sāmoan but not Pacific Islander: Tracing the ECLS-K Racial Data (Mis)Classification Journey

Wendy Castillo
Montclair State University

Daranee Taychachaiwongse Teng
University of Colorado, Denver

Kristine Jan Cruz Espinoza
California Lutheran University

Race is a socially and politically charged concept that remains contested in the United States. We examine racial data (mis)classification in the Early Childhood Longitudinal Studies (ECLS-K) dataset. Centering the racial data journey of Asian American, Native Hawaiian, and Pacific Islander (AA&NHPI) students, we find two types of racial data (mis)classification: (1) racial reformation related to the reconfiguration of parent/caregiver-reported racial data and (2) categorical friction when ethnicity was parent/caregiver-reported and race was not. Educational data practices and datasets like ECLS-K play a role in obscuring differentiated educational outcomes by operationalizing the myth that AA&NHPIs are a monolith. We offer recommendations for addressing racial data (mis)classification and engaging a critical race research praxis.

VERSION: May 2026

Suggested citation: Castillo, Wendy, Daranee Taychachaiwongse Teng, and Kristine Jan Cruz Espinoza. (2026). Hmong but not Asian, Sāmoan but not Pacific Islander: Tracing the ECLS-K Racial Data (Mis)Classification Journey. (EdWorkingPaper: 26-1486). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/wk9y-sm74

Hmong but not Asian, Sāmoan but not Pacific Islander: Tracing the ECLS-K Racial Data (Mis)Classification Journey

Wendy Castillo[1]
Montclair State University
1 Normal Avenue, Montclair, New Jersey 07043
castillow@montclair.edu

Daranee Taychachaiwongse Teng
University of Colorado, Denver
1201 Larimer Street, Denver, CO 80204
daranee.teng@ucdenver.edu

Kristine Jan Cruz Espinoza
California Lutheran University
60 W. Olsen Road, Thousand Oaks, CA 91360
kjespinoza@callutheran.edu

[1]Corresponding author

Abstract

Race is a socially and politically charged concept that remains contested in the United States. We examine racial data (mis)classification in the Early Childhood Longitudinal Studies (ECLS-K) dataset. Centering the racial data journey of Asian American, Native Hawaiian, and Pacific Islander (AA&NHPI) students, we find two types of racial data (mis)classification: (1) racial reformation related to the reconfiguration of parent/caregiver-reported racial data and (2) *categorical friction* when ethnicity was parent/caregiver-reported and race was not. Educational data practices and datasets like ECLS-K play a role in obscuring differentiated educational outcomes by operationalizing the myth that AA&NHPIs are a monolith. We offer recommendations for addressing racial data (mis)classification and engaging a critical race research praxis.

Keywords: Quant(Multi)Crit; race; racial categories; ECLS-K; racial reformation

The data that support the findings of this study are available at the National Center for Education Statistics at https://nces.ed.gov/ecls/kindergarten2011.asp. These data were derived from the public domain.

The authors used ChatGPT (GPT-5-turbo model) and Grammarly for grammar, spell-check, and language improvement when prompted for clarity.

Introduction

Race is a socially and politically charged concept that remains contested across academic, scientific, and public policy discussions in the United States. Racial data is characterized by socially constructed categories that are constantly redefined (Mills & Unsworth, 2018; Omi & Winant, 2014; Parker, 1998). While original notions of race were attributed to biology, ethnicity is "frequently assumed to be the cultural identity of a group within a nation state" (Grosfoguel, 2004, p. 315). Recognizing conventional distinctions between race and ethnicity, we use "racial data" both to encompass the broader set of constantly changing terminology, and as we orient to both race and ethnicity as socially-constructed concepts with material implications and consequences on the lived experiences of both "racialized ethnicities" and "ethnicized races" (Grosfoguel, 2004, p. 332). Although educational institutions and other governmental agencies collect racial data (among other demographic data) to monitor discrimination and inform resource allocation and policy decisions (Catalyst California, 2024; U.S. Census Bureau, n.d.), racial data integrity issues persist.

In this paper, we examine how the (mis)classification of racial data occurs among students in the Early Childhood Longitudinal Studies (ECLS-K) dataset. We use *(mis)classification* to refer to the classification practices and processes that reclassify racial data. This paper centers the racial data journey of Asian American, Native Hawaiian, and Pacific Islander[1] (AA&NHPI) students. At the same time, we include noteworthy patterns that surfaced in our analysis and findings relevant to other racially minoritized communities. Specifically, we ask: *How, and at what stage in the data journey, does (mis)classification of racial data occur for students in the ECLS-K dataset?*

From both educational and social policy and practice perspectives, Chang et al. (2015) assert that large federal data sets "have had limited applications for addressing the needs of [Asian American and Pacific Islander] communities" because they do not account for nuance that exists within the broad demographic label (p. 298). Further complicating these monolithic data practices, the challenges of Asian Americans (and Native Hawaiians and Pacific Islanders (NHPIs) by panethnic association) are often-overlooked in educational settings due to the *model minority* myth, which was founded on anti-Blackness and false notions of exceptionalism to drive a racist wedge between Asian Americans and other racially minoritized communities of color to perpetuate that Asians do not face problems (Hsieh & Kim, 2020; Lee, 1996; Poon et al., 2016; Suzuki, 1977; Yi et al., 2020).

In addition, critically interrogating racial data practices is especially salient now that more than one in 10 Americans identify with Two or more races in the U.S. Census and with 61% of all multiracial 3- to 5-year olds enrolled in school attending public schools (Jones et al., 2021; National Center for Education Statistics [NCES], 2022; U.S. Census Bureau, 2022). Following other scholars examining multiraciality in education (e.g., Cepeda & Combs, 2024), we use *multiracial* to name individuals identifying with Two or more racialized groups. We alternate use with *mixed race* if used in the original cited scholarship.

Consistent with the varying terminology used to describe individuals of more than one race, a cascade of changing categories and circumstances influence how multiracial individuals may (or may not) self-disclose and are counted (or not counted) in federal education datasets, with associated impacts on student identity development and learning. Based on their systematic review of studies on multiracial populations in education, psychology, sociology, social work, and public health, Charmaraman et al. (2014) describe racial categories as a "moving target" (p.

7). For one, multiracial individuals are uniquely racialized, with external, social perception influencing how they self-report monoracially (see Khanna, 2010), often resigning themselves to categories based on *reflected appraisal*, or how they think the world perceives them, rather than how they actually perceive themselves. Moreover, education researchers have found how multiracial identity development has ties to cognitive development, in which students are simultaneously negotiating multiple identities versus only one (Chaudhari & Pizzolato, 2008). As such, students can engage in developing multiracial critical consciousness to gain a deeper understanding of their multiple identities (Malaney-Brown, 2020).

At the same time, panethnic categories may not align with multiracial individuals' preferences. For instance, Asian and Latino populations often identify with their country of origin (e.g., Mexican, Dominican, Chinese, Filipino) over panethnic or regional labels (Latino, Southeast Asian; Lopez, 2013; Ruiz et al., 2023). Similarly, state-sanctioned categories may not accurately reflect how multiracial individuals refer to themselves, who instead use terms such as mixed race or biracial, or with specific labels such as AfroLatino, Blasian (Black and Asian), or Mexipina (Mexican and Filipina) (Baker et al., 2024). Regardless of how individuals self-identify, their preferred categories may not be available as options in surveys or forms. We argue: Even when individuals *do* select an option that represents them, (mis)classification occurs when individuals are reclassified and when data are later presented in reports.

Through (mis)classification, the racial data integrity becomes a pressing issue on two levels. First, reductive state-sanctioned categories flatten nuance and infringe on the dignity and self-determination of historically marginalized communities. Secondly, categories influence how inequities are documented and interpreted, thereby impacting policy and programmatic decisions, particularly for multiracial NHPI, Latino, and Black communities across educational

levels (early childhood, K-12, higher education). For example, the National Center for Education Statistics (NCES) annually calculates the allocation of federal funds for about $35 billion, of which $11 billion is allocated to Title I schools (NCES, n.d.). As a postsecondary education example, the racial enrollment count of monoracial AA&NHPI undergraduates is used, in part, to determine eligibility for the federal Asian American and Native American Pacific Islander-Serving Institution designation and competitive grant program (Espinoza et al., 2025).

This descriptive study was prompted when Author 1 noticed inconsistent racial data patterns when using the ECLS-K dataset for a separate project. As a [redacted for peer review] scholar whose work focuses on applying a racial equity lens to quantitative methods, Author 1 invited Authors 2 and 3 (who respectively identify as [redacted for peer review] and [redacted for peer review]) as research collaborators to investigate the incongruent patterns in the data. Author 2 engages global geopolitical and historical contexts to explore possibilities for liberatory knowledge production in quantitative research through a decolonial and anti-imperialist lens. Author 3 has been studying classifications in education, including racial data, through a critical lens in pursuit of racial equity. In all, rather than including a positionality statement with "backward-looking identity narratives aimed primarily at affirming the researcher's legitimacy in knowledge production" (Zembylas, 2025), we ground this work in our shared histories and transnational contexts impacted by western colonialism and imperialism to inform our responsibility as researchers to transform the data narratives and practices that have had and continue to have harmful consequences for the lived experiences of communities of color.

Our study contributes new insights to critical quantitative scholarship by addressing the following underexamined areas of educational research: (a) uncovering the stages in the data process where inconsistencies appear between parent/caregiver-reported racial data and the final

ECLS-K reporting, particularly for multiracial students, and (b) surfacing how already small ethnic subgroups are drastically undercounted while other groups maintain substantive racial data representation. This study traces when and how racial (mis)classification happens, while also surfacing an important ethical imperative for researchers working with commonly used large datasets. Indeed, racial data (mis)classification likely persists beyond ECLS-K. As such, researchers must engage in an ongoing critical research praxis enabling them to recognize data discrepancies, particularly when such inconsistencies pose harm to historically marginalized communities. We offer recommendations and analytical guidance to mitigate racial data integrity issues discussed in this paper.

Theoretical Frameworks

We use racial reformation and Quant(Multi)Crit to inform our understanding of racial data (mis)classification processes and practices in relation to the ECLS-K dataset. Racial reformation builds on Omi and Winant's (2014) *racial formation*, originally defined as "the process by which social, economic and political forces determine the content and importance of racial categories, and by which they are in turn shaped by racial meanings" (p.10). Coined by Campbell-Montalvo (2020) and situated within an educational context, *racial reformation* describes the alteration of an individual's racial/ethnic self-identification during data processing and final reporting. Specifically, racial reformation results from human error when K-12 school staff force the operationalization of racial and ethnic groupings. In this forced operationalization, some students may be subject to (mis)classification, which in turn leads to final reporting that may not align with originally selected race and ethnicity options. In this paper, we use racial reformation to frame our examination around how data processing decisions in the ECLS-K

dataset can lead to discrepancies between parents/caregivers' racial identification of their multiracial children and the final reporting of their children's race and ethnicity.

Secondly, we draw on Quant(Multi)Crit to offer an additional layer of nuance regarding the biases and problematic data practices that impact multiracial individuals. Placing Critical Multiracial Theory (MultiCrit; Harris, 2016) in conversation with Quantitative Critical Race Theory (QuantCrit; Garcia et al., 2018; Gillborn et al., 2018), Quant(Multi)Crit (Vezaldenos, 2025) extends Quant(Crit) to attend to monoracial structures and systems. Of the expanded Quant(Multi)Crit tenets advanced by Vezaldenos (2025), we leverage the following in our analysis to surface racial reformation patterns in the ECLS-K dataset: (a) the non-neutrality of numbers in perpetuating monoracist racial categories, (b) that the generic multiracial data category flattens nuance and diversity by grouping all mixed race individuals together as part of a monolith, and (c) using numbers for social justice rather than enabling the continuation of monoracism in quantitative research. Appendix A further details Quant(Multi)Crit in comparison to QuantCrit, including corresponding definitions and tenets.

Background

Education data systems use the U.S. Office of Management and Budget's (OMB) minimum reporting categories, which include "Hispanic, race specified" and "Hispanic, no race specified" ethnicity options as well as six race categories: "American Indian or Alaska Native, " "Asian," "Black or African American," "Native Hawaiian or Other Pacific Islander," "Two or more races," and white[2]). Nonetheless, a racial data integrity issue persists when individuals are forced into these categories, enabling the (mis)classification of student racial data. Elsewhere, inconsistencies between self-reported racial data compared to institutional records have been extensively studied in public health research with empirical studies consistently reporting a

disproportionate impact on communities of color (Azar et al., 2012; Hahn et al., 1992; Zaslavsky et al., 2012). On the other hand, the extent to which racial (mis)classification, or the reconfiguration processes of racial data, in educational contexts remains sparse. Recognizing the interconnected nature of public health and education, which share common methodological approaches to research on social determinants of patient or student outcomes, insights from public health scholarship on racial data integrity inform this paper.

Concordance Between Self-Reported Racial Data and Administrative Records

Discrepancies between students' self-reported and observer-reported racial data across the United States and its colonial territory of Puerto Rico have been found in educational studies since at least over 40 years ago (Ginorio & Berry, 1972; Massey, 1980; Rivera & Pennock-Roman, 1987; Schneider, 1980). Additional research on inconsistencies between self-identified racial data and administrative records has remained sporadic and sparse in the field of education. Withstanding the examples presented in this section, to our knowledge, racial reformation in K-12 and higher education has not been widely documented. The few studies in education are concerning, given the pervasiveness of problematic racial data collection practices across educational levels, including K-12 institutions, which threaten the validity of student demographic data (Ford, 2019).

Campbell-Montalvo (2020, 2021) documented how public schools in Florida changed original racial data to fit the requirements of the Florida Department of Education (FLDOE). Following OMB guidelines, the FLDOE's reporting requirements consist of six race categories and two Hispanic ethnicity categories (race specified and no race specified), and state law dictates that ethnicity supersedes race. Moreover, students are not allowed to select both a racial group and an ethnicity, with non-Latinos in more than one racial group coded as "multiracial"

(Florida Legislature, 2023). Using a qualitative observational method, Campbell-Montalvo (2020; 2021) examined how white monolingual school personnel reconciled raw race and ethnicity data before sending them to the FLDOE. When Campbell-Montalvo observed that a family selected a race *and* ethnicity in the school form, she asked school personnel to discuss their reasoning for the final report determination, which unveiled their biases and assumptions. For example, a school staff member used her knowledge of the family composition, explaining, "Dad is Black, mom is Hispanic, so I'd go with Black" (Campbell-Montalvo, 2020, p. 188). When families selected white and another race, school personnel defaulted to white, potentially artificially inflating the white student population in this case from 30 to 33 percent.

Concerns about the harms of observer classification have been raised for decades. Over 15 years ago, in a letter to the Washington State Office of the Superintendent of Public Instruction (OSPI), the superintendent of a public school district in the Seattle metropolitan area challenged a state mandate for school districts to guess the race and ethnicity of students whose demographic data were missing from enrollment and registration paperwork:

> Perhaps you or someone else at OSPI or (U.S. Department of Education) could help me understand how I train my staff to visually recognize the differences between a Hmong or Vietnamese, or between a Fijian and a Samoan, or between a Lummi and Makah. [...] In the era when all school districts are struggling mightily to make our organizations open and welcoming to the vast variety of immigrants entering each year, we are now required to simply guess at who they are, what they are, from whence they came. There may be no more insulting an act that we could perform for incoming families (Low, 2010, para. 13).

The school district leader's statement reveals equity implications. Educators and practitioners have a responsibility to question legislative mandates and policy directives that infringe upon the dignity of historically marginalized students and families.

Racial (mis)classification consequences have also been studied at the postsecondary level as related to the U.S. Department of Education's NCES Integrated Postsecondary Education Data System (IPEDS) guidelines (Espinoza et al., 2025; Wong-Campbell & Ramrakhiani, 2024). Espinoza et al. (2025) analyzed administrative data on one institution's undergraduate population to examine the differential proportions of racially minoritized students following IPEDS guidelines. Namely, based on IPEDS racial classification guidelines, Espinoza et al. (2025) found that this led to the recoding and (under)counting of NHPI, American Indian or Alaskan Native (AIAN), Asian, and Black or African American undergraduates as being grouped into categories as Two or more races or Hispanic or Latino, with significant consequences for NHPI and AIAN undergraduates. Wong-Campbell & Ramrakhiani (2024) similarly examined the consequences of multiracial (re)categorization because of IPEDS guidelines. Based on their institutional sample of student records, Wong-Campbell & Ramrakhiani (2024) calculated that self-labeled multiracial college students were more likely to be recategorized per IPEDS guidelines, with self-identified AIAN and NHPI students as having higher odds of recategorization.

Transnational Context for State-Sanctioned Categories: The Legacy of U.S. Imperialism

We situate our study around how AA&NHPIs, including multiracial AA&NHPIs are represented in data, within the transnational contexts of both historical and contemporary U.S. empire. A growing body of scholarship details how positioning Asian Americans (and NHPIs by panethnic association) as a model minority obfuscates and conceals imperial legacies of U.S.

military conquest on the Asian continent and across the Pacific Islands while obscuring the complex struggles encountered by many members within the panethnic AA&NHPI category (Bascara, 2006; Parker & Taychachaiwongse Teng, 2024; *Rice v. Cayetano*, 2000). Of particular importance are the often over-looked nuances of the varying political statuses of Pacific Islanders under U.S. colonialism, especially related to military occupation in the Pacific (i.e., U.S. citizens, Compact of Free Association migrants, U.S. nationals, immigrants without formal affiliation; EPIC, n.d.).

We trace the historical encounters of Asia and the Pacific Islands with the U.S. empire to illuminate the geopolitical underpinnings of intermarriage that has shaped the present-day realities of multiracial AA&NHPIs. For instance, the emergence of Asian war brides–legitimized by the War Brides Act of 1945–grew from contact between American soldiers and Asian women during World War II. Pacific Islander war brides also surfaced as a sociological phenomenon resulting from U.S. occupation of the Cook Islands, Fiji, Tonga, and Western Sāmoa during World War II (Wanhalla, 2023). U.S. military presence in Korea and Vietnam perpetuated the social construction of war brides (Saenz, 1994; Zeiger, 2010). In *War Baby/Love Child*, an interdisciplinary anthology, Critical Mixed Race and Asian American Studies scholars explored implications of the Immigration and Nationality Act of 1965 that expanded opportunities for Asian immigration against the backdrop of historical legacies of American imperialism in Asia and the Pacific Islands: "[T]he mixed Asian body writ large continues to signify specific histories of Asian Pacific-U.S. collisions: narratives of war, economic and political migration, and colonization" (Kina & Dariotis, 2012, p. 5). Indeed, a historical analysis of U.S. foreign and domestic policy offers a nuanced grounding of the transnational contexts of American

imperialism that influenced the present-day racialization of mixed race AA&NHPIs, including AA&NHPI (mis)classification in education research.

Methodology

Researchers engaging in primary data collection entails gathering data firsthand. Alternatively, secondary data are data collected by someone else. This however, does not mean that data users and researchers simply accept the secondary data *as is*. From a Quant(Multi)Crit perspective, researchers have a responsibility to interrogate the theoretical principles that guided the research process to identify discrepancies in quantitative data collection, analysis, and reporting practices related to secondary datasets (Vezaldenos, 2025).

This descriptive study explores racial data within the publicly available ECLS-K dataset, which followed a cohort of kindergarteners through fifth grade from 2011 through 2016. The dataset includes a sample of over 18,000 students' demographic and academic information. Although this study focuses on AA&NHPI students because of initial observations, we include our observations of the broader impact of (mis)classification occurring for all racialized groups in the dataset.

We examine the distribution of students within the ECLS-K dataset by comparing the frequency (*n*) and percentage (%) across all available race and ethnicity variables: 1) a composite variable of parent-report race and ethnicity variable, 2) a composite variable merging race collected via parent-report with school records, 3) binary race variables, and 4) specific ethnicity variables. When reporting data as they appear in ECLS-K, we use the word *parent*. When otherwise referring to the adults who reported racial data for the students represented in the dataset, we use the term *parent/caregiver*. Rather than treating parents and caregivers as interchangeable terminology, we acknowledge the legal, societal, and political presumptions in

the ECLS-K that privilege an imposed legitimacy of biological parents in heteronormative marriages over diverse family structures and non-parent caregivers, who are "disproportionately populated by persons of color" (Coupet, 2010, p. 595; Fineman, 2006; NCES, 2011).

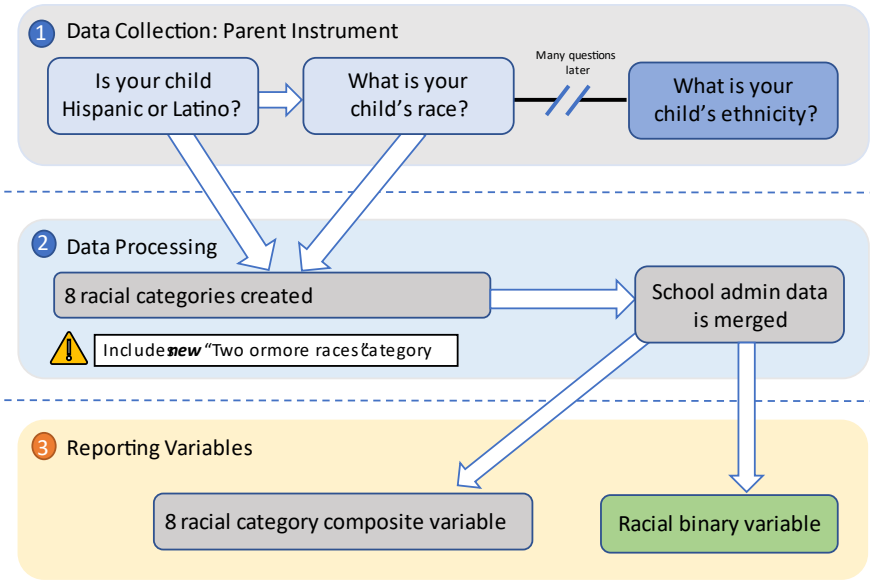
By evaluating alignment across the four categorical variables, patterns of systematic shifts in how students are represented in the racial data emerged. We focused the majority of our analysis on the second variable (Table 1 column 2) using it as our reference point. Notably, this variable in the ECLS-K dataset has the most complete data and is often most widely used by data analysts due to its larger sample size.

Results

We examined patterns of racial data transformation of students in the ECLS-K dataset from the data collection to the reporting phase. Data collection includes racial data reported by parents/caregivers and school administrators. By reporting phase, we refer to the final cleaned administrative data reported by ECLS-K that is publicly available and ready for a researcher to analyze, using one of multiple racial variables in the dataset. We traced the data journey of the ECLS-K racial data in Figure 1.

Figure 1

Tracing the ECLS-K Racial Data (Mis)Classification Journey



Note. Racial data (mis)classification occurs when multiracial students are collapsed into a "Two or more races" category created by NCES. Racial data (mis)classification also occurs when ethnicity data reported by parents/caregivers are ignored because the parent only reported an ethnicity category, but not race category for their child.

Data Collection Stage

The ECLS-K data journey begins with data collection as shown in Figure 1. In the ECLS-K instrument, the parent/caregiver is first asked whether their child is Hispanic or Latino (see Appendix B). They are subsequently asked their child's race, allowing them to select one or more of the following race categories: white, Black or African American, Asian, Native Hawaiian or Other Pacific Islander, and American Indian or Alaskan Native.

In addition to the questions about Hispanic or Latino ethnicity and race, parents/caregivers are asked about their child's ethnicity in a separate question, located much later in the instrument. The exact wording of the question states: "While we asked about this in the past, I have some additional questions about {CHILD}'s ethnicity ..." "Which of the following Asian groups best describes {CHILD}'s origin? Asian Indian, Chinese, Filipino,

Japanese, Korean, Vietnamese, Hmong, or Other Asian? (mark all that apply)," "Which of the following Pacific Islander groups best describes {CHILD}'s origin? Native Hawaiian, Guamanian or Chamorro[3], Samoan, or Other Pacific Islander? (mark all that apply.)" and "Which of the following Hispanic groups best describes {CHILD}'s origin? Cuban, Mexican, Puerto Rican, or Other Hispanic? (mark all that apply.)" There were no follow-up questions for AIAN asking about tribal affiliation.

Data Processing Stage

For ECKL-S, the U.S. Department of Education data analysts process data collected from the schools to accordingly match with the eight minimum categories set forth by OMB (NCES 2011). This processed data is later provided for researchers and analysts to use.

1. white, non-Hispanic
2. Black non-Hispanic
3. Hispanic, race-specified
4. Hispanic, no race specified
5. Asian, non-Hispanic
6. Native Hawaiian or Other Pacific Islander
7. American Indian or Alaskan Native, non-Hispanic
8. Two or more races, non-Hispanic

The eighth category represents children whose parent/caregiver selected multiple racial categories, but were assigned to a newly created Two or more races category by the data analysts working under NCES . In the ECLS-K, as with other federal datasets, a separate multiracial reporting category is not presented as an option. Instead, individuals may select multiple options and, if they do, are placed into the Two or more races category. If parents/caregivers select

Hispanic or Latino for their student's ethnicity, that choice supersedes any race selection (including selecting more than one race).

Next, school administrative data is merged with parent/caregiver-report data to create a second composite variable that also includes the Two or more races category. Nonetheless, priority is given to the original selection the parent/caregiver made on ECLS-K questionnaires. In other words, school administrative data is only used for students in the dataset with missing or unknown racial data. According to the ECLS-K manual, administrative data is "collected about the study schools, school staff, and children from available administrative records or existing data sources (such as the Common Core of Data) or from conversations between data collection staff and school staff" (NCES, 2019, p. 7-1). Finally, for the ethnicity variable, there are no data transformations that occur during the processing stage.

Data Reporting Stage

Those using the publicly available and restricted ECLS-K data may choose from four distinct variables: (a) composite parent reporting with the addition of the Two or more races category by NCES; (b) composite parent and school administrator reporting; (c) binary racial data based on parent and school administrator reporting; and (d) parent-reported detailed ethnic subgroup data.

Variable 1: Composite Parent-Reported Race and Hispanic/Latino Ethnicity (with NCES-Created Two or more Races Category)

Column 1 in Table 1 shows the first composite race variable, which represents both original and manipulated data, including: (a) data collected via the parent interview instrument for a standalone Hispanic or Latino ethnicity question and a separate question with five race

category options, and (b) data reflecting a newly-created Two or more races category. This variable represents the racial data of almost 17,000 students.

Table 1

Racial Composite Variables in the ECLS-K:2011 Dataset

	1. Composite Parent Interview		*2. Composite Parent*
Race/Ethnicity Category	*n*	*%*	*n*
Asian, non-Hispanic	1,410	8.31	1,543
American Indian or Alaska Native, non-Hispanic	140	0.83	168
Black/African American, non-Hispanic	2,152	12.68	2,396
Hispanic, race specified	4,192	24.71	4,207
Hispanic, no race specified	132	0.78	385
Native Hawaiian/Other Pacific Islander, non-Hispanic	97	0.57	117
White, non-Hispanic	8,026	47.30	8,488
Two or more races, non-Hispanic	819	4.83	827
Total	16,968	100.00	18,131

Variable 2: Composite Parent- and School Administrator-Reported Race and Hispanic/Latino Ethnicity (with NCES-Created Two or more Races Category)

The second variable (seen in column 2 in Table 1) merges the racial data represented in the first variable with school administrative data, and is typically the most widely used because it

has more complete data and thus a higher sample size. Column 2 shows available data for more than 18,000 students, most of whom are white. However, the methods used to collect the additional school data are unknown, limiting our ability to assess their reliability and validity.

As mentioned previously, the Hispanic or Latino selection superseded any racial selection. Thus, students with both Hispanic or Latino ethnicity and any other race categories selected during the parent interview were counted only as Hispanic or Latino, but were not counted in their respective chosen racial category/categories. Table 2 shows that there were a total of 275 Hispanic/Latino students with a specific race selected (who were not white) who were not counted in the counts of Table 1.

Table 2

Hispanic/Latino Students with a Race Specified Who Were Not Counted in Their Selected Race Category Because Hispanic/Latino Supersedes Any Race Selection

Race/Ethnicity Category	*Hispanic/Latino students*
	n
Asian	36
American Indian or Alaska Native	54
Black	157
Native Hawaiian/Other Pacific Islander	28
Total	275

Variable 3: Binary Parent- and School Administrator-Reported Race and Hispanic/Latino Ethnicity (with NCES-Created Two or more Races Category)

Based on the combined responses from the parent interview instrument and administrative data, the ECLS-K data team developed a third variable, binary race (e.g., Black/Not Black, Asian/Not Asian), shown in Table 3. This variable allows students to be counted in multiple racial categories. For example, if parents and/or school administrators indicate that a student is Black and Asian, then the same student is counted both in the Black category *and* Asian category. In addition, the student is also counted in the Two or more race binary category. Similarly, if a parent/caregiver selects Hispanic or Latino along with a race category, the student is counted in both groups. However, the binary variable does not take ethnicity categories beyond Hispanic or Latino into account.

Of note, when using this binary variable, there are large increases in the overall number of students of color since it allows for duplicated student counts, which more comprehensively reflects complex identities compared to the previously discussed variables. While representing the two smallest numerical increases, the largest proportional increases are for AIAN (+155.9%) and NHPI (+112.8%) compared to the commonly used composite race variable in Table 1, column 2. These differences reflect the reallocation of students who were previously classified under Two or more races in Table 1, and are now counted within each of their selected race categories. A multiracial variable was created for students whose parents/caregivers selected Two or more races. It is unclear why the category name differed from previous variables, yet it is aligned with other naming inconsistencies in the ECLS-K Data (e.g. using “Black” and “Black or African American” interchangeably). Nonetheless, the parent interview guide does not provide an option to select a category called “Multiracial.”

Table 3

Binary Race Variable in the ECLS-K:2011 Dataset

Race/Ethnicity Category	Binary Race/Ethnicity Indicators (duplicated data)		Difference from column 2 in Table 1	
	n	*%*	*n*	*%*
Asian	1,765	10.48	355	+14.4
American Indian or Alaska Native	430	2.55	290	+155.9
Black	2,840	16.87	688	+18.5
Hispanic or Latino	4,324[a]	25.47[a]	—	—
Native Hawaiian/Other Pacific Islander	249	1.48	152	+112.8
White	12,688	75.36	4662	+49.5
Multiracial	1,038	6.16	219	+25.5
Total	—-	—-[b]	6366	—

[a] Hispanic/Latino (any race) is a combined Hispanic category representing individuals regardless of whether race was specified. The Hispanic/Latino (any race) category appeared when binary race/ethnicity variables were used in the ECLS-K dataset.

[b] As these categories are not mutually exclusive, the column labels (%) will not add up to 100%.

Variable 4: Parent-Reported Detailed Ethnic Subgroups

Ethnicity data is the fourth variable option that can be used when requesting ECLS-K data. Interestingly, the codebook does not mention this particular variable in the race and ethnicity section (see NCES, 2019, pp. 7-20). Therefore, to obtain this reported data, you must either *a priori* know the ethnicity variable exists or look through the 219-page instrument and find it. In Table 4, column 1, the data shows that data collected during the parent interview identified 110 children as NHPI, 1199 as Asian, and 2864 as Hispanic/Latino. However, this data does not align with the reported race variables in Table 1.

Data inconsistencies were significant for NHPI students. When asked "Which of the following Pacific Islander groups best describes {CHILD}'s origin?" there were 10 students whose parents/caregivers selected Native Hawaiian as their child's ethnicity; however, they did not select NHPI in the preceding race category question in Table 1 column 2. Similarly, 28 children were marked as Other Pacific Islander" for the ethnicity question while NHPI was not selected in the race category question. There were two students whose race was not reported as NHPI, but one was marked only as Sāmoan and one as only Guamanian/Chamorro. In all, while only 41 students had only an NHPI ethnic subgroup (but not the NHPI race category) selected, this proportionally represented a 37.3% sample loss from the total number of NHPI students reported (see column 2 in Table 1). This same phenomenon happened with students who may have otherwise been classified and counted in the aggregate as Asian (n=121) or Hispanic/Latino (n=133), though of a lesser proportional sample loss (7.8% difference for Asian and 2.9% difference for Hispanic/Latino). See columns 2 and 3 of Table 4 for more details.

Table 4

(Mis)classification Between Composite and Binary Race Indicators and Detailed Ethnicities

Race Category and Corresponding Ethnicity	Parent/ Caregiver-reported	Selected Ethnic Subgroup, but Did *Not* Select Corresponding Race Category	Difference from Table 1 column 2
	n	*n*	*%*
NHPI			37.3%
Hawaiian	35	10	
Guamanian/Chamorro	8	1	
Samoan	19	2	
Other Pacific Islander	48	28	
NHPI Total	110	41	
Asian			7.8%
Asian Indian	209	14	
Chinese	364	13	
Filipino	182	51	
Hmong	11	1	
Japanese	79	12	
Korean	87	7	
Vietnamese	149	5	

	Other Asian	118	18	
	Asian Total	1199	121	
Hispanic/Latino				2.9%
	Cuban	71	7	
	Mexican	2049	66	
	Puerto Rican	176	14	
	Other Hispanic	568	46	
Total	Hispanic/Latino	2864	133	

Note: The ECLS-K user manual and codebook did not provide a rationale for how or why these ethnicity categories were options provided on the parent data collection instrument.

Discussion

Our findings revealed two distinct types of racial data (mis)classification occurring in the ECLS-K dataset. The first type is racial reformation related to the reconfiguration of parent/caregiver-reported racial data through institutional data processing and reporting practices. While Campbell-Montalvo (2020), uncovered racial reformation occurring at the school level, our findings reveal the same phenomena happening at the level of federal institutions. Racial reformation occurred across the data collection, processing, and reporting stages of the data journey. The second form of (mis)classification is a particularly salient discovery that we refer to as *categorical friction*, which occurred when ethnicity was parent/caregiver-reported and race was not. Categorical friction was identified during the data collection stage resulting in undercounting of specific subgroups as manifested during the

reporting stage of the racial data journey. Both racial reformation and categorical friction are detailed in this section as two related, but distinct racial data (mis)classification phenomena.

Racial Reformation in ECLS-K

Racial reformation occurs in the ECLS-K in two distinct ways: (a) when a parent/caregiver selects more than one race for their child's identification, their child is then assigned by NCES to a monolithic Two or more races category, and (b) when the Hispanic/Latino ethnicity selection overrides the race(s) reported by parents/caregivers. Racial reformation was observed for the first two reporting racial variables in ECLS-K: (1) composite parent reporting and (2) composite parent and school administrator reporting. Racial reformation resulted in a decrease in the number of students counted in the parents/caregivers' originally selected categories. For both of these composite racial variables, the detailed ethnicity categories that appear much later in the ECLS-K instrument are completely ignored and not included in the final ECLS-K race and ethnicity reporting section. Specifically, for the second race variable, when administrative school data is added to parent/caregiver-reported racial data, we are unsure whether any *additional* racial reformation also happens because of the varying sources of school data.

When examining how the binary race variable was created, racial reformation did not necessarily occur. Although all originally selected categories were kept intact, a multiracial category was applied to students with more than one race indicated, despite the fact that parents/caregivers never selected multiracial as their response. In addition, similar to the previous racial variables, ethnicity is completely ignored. While the binary variable supports more expansive applications of race data, its use in relation to ethnicity data rules must be considered to ensure accurate representation of students in the dataset. When the racial data

practices by the ECLS-K data team at NCES are applied (i.e. grouping those who selected more than one race into one category and the Hispanic/Latino selection superseding any race selection), there is a substantial loss in Asian, Black, NHPI, and AIAN students whose parents/caregivers identified their children as more than one race (*n*=1485 or 8.2% of the total sample). The greatest impact of racial reformation is on populations with smaller sample sizes, such as NHPI and AIAN students. As mentioned previously, the second race composite variable is our reference point (see Table 1, column 2 for reference). In addition to the sample loss that occurred for Asian, AIAN, Black, NHPI, and multiracial students, it is also possible that racial reformation happened through parents/caregivers' reporting. For instance, students may identify differently from the categories their parent(s)/caregiver(s) selected. However, it is difficult to validate or quantify such errors when we do not know about the data collection and processes from the respective school districts participating in the ECLS-K. Ultimately, the racial (mis)classification in the ECLS-K provides a tangible example of pervasive monoracial normativity (Vezaldenos, 2025) and, ergo, the Quant(Multi)Crit tenets of monoracism and numbers upholding monoracist racial categories.

Categorical Friction in ECLS-K

For the specific variables that asked parents/caregivers to report ethnicity, a different form of (mis)classification emerged: categorical friction. In this paper, we define *categorical friction* as the misalignment between the self-identification of racialized communities and the reductive state-imposed categories used in federal data reporting. In the case of ECLS-K data, categorical friction leads to undercounting of students whose ethnicity was selected by the parent/caregiver, but *not* the race category that the ethnic subgroup is typically nested within. We report this occurrence for 133 students. Although categorical friction does not impact a large

proportion of the total sample, for numerically smaller populations like NHPIs, an additional 41 students would increase the NHPI student sample by 37%. The incidence of a subset of parents/caregivers selecting their child's ethnicity on the questionnaire while *not* reporting race could be viewed as a reliability issue simply due to respondent oversight; however, the systematic nature of this occurrence suggests the presence of a more substantive sociological phenomena.

Prevailing views on race "presumes that racial categories assigned by the state align or correspond with actual group identities and behavior" (Okamoto, 2014, p. 3). Moreover, Junn & Masuoka (2008) assert that definitions of race in the United States are "uniquely American constructs" and that immigrants "may not adhere to the categories imposed upon them," particularly regarding the broader racial labels their ethnicities are grouped within by state and federal reporting standards (p. 730). Providing ethnicity data, but *not* the race data according to state-imposed categories (e.g., selecting Sāmoan but not NHPI), a phenomenon we call *categorical friction,* suggests a form of resistance to the construction of race in the United States. Conversely, some AA&NHPI parents/caregivers may be unfamiliar with race constructions in the United States (Espiritu, 1992) and/or may feel a tension similar to Khanna's (2010) concept of reflected appraisal, forcing them to select the race ascribed onto them. Okamoto (2014) notes that members of distinct ethnic groups, "by taking part in creating their own collective histories, cultures, and identities, are challenging and redefining current notions of race, despite the fact that group boundaries are often set by the larger society" (p. 24). Categorical friction is a phenomenon that underscores the socially (and politically) constructed nature of race and ethnicity categories, leading to the (under)counting of Asian, NHPI, and Hispanic/Latino students. In turn, these counts bear implications for equitable distribution of resources and the

dignity and self-determination of communities who are already small in numbers in U.S. educational contexts, often rendering them invisible in quantitative datasets. Although the data collection displayed in Table 4 is aligned with Quant(Multi)Crit recommendations that mitigate racial reformation in the data, it does not preclude other types of (mis)classification from occurring. Our findings surface a pressing need for researchers using large secondary datasets to critically engage with the possibility of data incongruence with implications for historically marginalized communities.

Conclusion and Recommendations

Present-day educational data practices and datasets like ECLS-K continue to operationalize the myth that AA&NHPIs are a monolith, which plays a role in invisibilizing or obscuring the disparate educational and social outcomes within the AA&NHPI community. Although the inherent choice of using categories associated with race and ethnicity will likely always "obscure the complex, heterogenous, and contradictory relationships between and within such groupings" (Grosfoguel, 2004, p. 319), this does not absolve the researcher or data user from questioning and understanding how these categories were formed, who chose them, and how to use them moving forward. Here, we interrogated and analyzed the ECLS-K dataset to identify how, and at which stages during the data journey racial reformation and categorical friction occurred.

Finally, a noteworthy observation that surfaced throughout our analyses was inconsistent racial data terminology found across the ECLS-K resources, including the publicly available codebook, parent interview instrument, and the data file. For example, students with more than one race category selected by their parent/caregiver were sometimes referred to as Two or more races or multiracial at other times, but no glossary of terms was offered. Another example was

the use of Black as a race category on certain pages of the codebook, while Black or African American was also used to refer to the same category (in both the codebook and the parent interview guide) without a rationale for or acknowledgement of the interchangeable term use. We observed a similar inconsistent use of Hispanic alone in the dataset in addition to the use of Hispanic or Latino across the ECLS-K written resources. Throughout our reflexive research process, we noticed how the inconsistencies in the terminology used to describe racial data often created confusion for ourselves as researchers. Indeed, our observations here speaks to other scholars' call for more careful reflection on (and deeper engagement with) racial data terminology used in empirical research and writing (Baker et al., 2024). Such observations pushed us to meticulously identify and interrogate the incongruences, ultimately affirming the Quant(Multi)Crit and QuantCrit presuppositions that categories (especially multiracial categories) are neither neutral nor natural.

In what follows, we offer two sets of recommendations: (a) analytical guidance for addressing the racial data (mis)classification, and (b) ways to engage a critical race research praxis.

Analytical Guidance for Addressing Racial Data (Mis)classification

1. **Verify whether the racial data questionnaire or survey allowed more than one selection.** If more than one selection was allowed, investigate the final reporting of individuals to determine whether racial reformation occurred.
2. **Check if Hispanic or Latino was presented as a separate ethnicity question.** If so, identify whether ethnicity was merged with race *and* if the Hispanic or Latino category supersedes any racial selection or not.

3. When available, use **binary race and/or ethnicity variables**. For instance, binary race variables would allow for students counted in both Hispanic or Latino *and* Asian to be counted. Note that these binary variables represent duplicate counts.
4. **Check if other race or ethnicity questions were asked**. While additional options may not be the *de facto* variables to use does not mean they should be ignored.
5. **Identify inconsistencies in racial terminology** in large datasets and their accompanying materials (codebook, instruments). Note incongruences and provide a consistent rationale for how you decide to use specific naming conventions or inclusion criteria in your own analyses and reporting (Baker et al., 2024).

Engaging in Critical Race Research Praxis

1. **Treat theory and method as inseparable**. Integrate critical theoretical principles throughout the research process to identify discrepancies in quantitative data collection, analysis, and reporting practices (Castillo & Strunk, 2025; Strunk, 2025). This applies with secondary data, as we did in this paper.
2. **Situate research within historical and transnational contexts.** A foundational premise of CRT, which ultimately undergirds Quant(Multi)Crit, demands historical context. Furthermore, to challenge impulses of ahistoricism that insulate educational research within a United States frame, interrogate how historical and ongoing imperialism and military conquest across the globe have shaped understandings of racialization and social science research more broadly (Ake, 1991).
3. **Navigate the nuances of data disaggregation.** Disaggregating racial data in educational research reveals varied educational outcomes of AA&NHPI students (CARE, 2011; Nguyen et al., 2017). However, these disaggregated data are not immune to gap-gazing,

the tendency to center a more dominant group in data comparisons from a deficit perspective (Gutiérrez & Dixon-Román, 2011) as disaggregation can unintentionally lead to the reification of negative stereotypes amongst AA&NHPI subgroups such as Southeast Asian and East Asian, or NHPIs compared to Asian Americans (Poon, 2017). Moreover, it is important to consider varying levels of disaggregation in order to surface clearer insights specific to the unique contexts of various subgroups, such as distinctions between Asians and Pacific Islanders, within the Pacific Islander label, and within the Two or more races category.

Drawing on Kānaka Maoli (Native Hawaiian) wisdom, we urge researchers and practitioners using institutionally-produced datasets at any level (e.g., district, state, federal) to *e maka'ala*: to remain alert, vigilant, and mindful to humanize data and honor our commitments to liberatory scholarship (Brandehoff, 2025; Watkins-Victorino, Native Hawaiian, research director at the Office of Hawaiian Affairs, personal communication, January 15, 2026). We argue that the greatest threat to data reliability and validity in social science research is human error, with the accuracy of demographic estimates also dependent on the race and ethnicity data sources (Boehmer et al., 2002; Cooper et al., 2021; Sojka et al., 2025). Grounding in critical theoretical principles alongside broader historical and transnational contexts will support the identification of counterintuitive patterns in commonly-used datasets, enabling scholars to disrupt harmful assumptions and data practices, including the (mis)classification of racial data, that compromises the dignity of historically marginalized communities and the integrity of educational research.

Notes

1. It is important to note that various panethnic labels have been used to describe Asian and Pasifika communities in the United States, including *Asian American and Pacific Islander* (AAPI), *Asian Pacific American* (APA), and *Asian Pacific Islander Desi American* (APIDA). We use AA&NHPI to refer to this panethnic community, one that has since been divided into separate Asian and Native Hawaiian and Other Pacific Islander (NHPI) categories in federal racial data reporting, yet is still often treated as a single aggregated group. While an extensive discussion of AA&NHPI panethnicity is beyond the scope of this study, Gogue et al. (2022) and Omi et al. (2020) offer more robust analyses of panethnic racial formation.
2. Engaging the QuantCrit tenet of the *centrality of racism*, we intentionally lowercase white to de-center whiteness, foregrounding the importance of engaging a critical race lens when interrogating data and data practices that perpetuate racism.
3. The term *Chamorro* is used in this paper as it appears in the ECLS-K dataset; however, we acknowledge the contested linguistic and orthographical distinctions between the use of *Chamorro* in the Northern Mariana Islands and *CHamoru* in Guam, two territories that were colonized by Spain, Germany, Japan, and currently exist under United States' colonial rule (Madrid & Cepeda, 2019; Taitano, n. d.).

References

Ake, C. (1991). *Social science as imperialism: The theory of political development.* Ibadan University Press.

Arab American Institute. (2024, March 28). *Arab American Institute welcomes new "Middle Eastern or North African" (MENA) category and revision of race and ethnicity standards: Hailed historic category regrettably fails to capture full diversity of Arab American community* [Press release]. https://yallacountmenain.org/s/PRESS-RELEASE-spd15.pdf

Azar, K. M., Moreno, M. R., Wong, E. C., Shin, J. J., Soto, C., & Palaniappan, L. P. (2012). Accuracy of data entry of patient race/ethnicity/ancestry and preferred spoken language in an ambulatory care setting. *Health Services Research*, *47*(1 Pt 1), 228–240. https://doi.org/10.1111/j.1475-6773.2011.01305.x

Baker, D. J., Ford, K. S., Viano, S., & Johnston-Guerrero, M. P. (2024). Racial category usage in education research: Examining the publications from AERA journals. *AERA Open, 10*(1), 1-19. https://doi.org/10.1177/23328584241293681

Bascara, V. (2006). *Model-Minority Imperialism*. University of Minnesota Press.

Boehmer, U., Kressin, N. R., Berlowitz, D. R., Christiansen, C. L., Kazis, L. E., & Jones, J. A. (2002). Self-reported vs administrative race/ethnicity data and study results. *American Journal of Public Health, 92*(9),1471–1472. https://doi.org/10.2105/ajph.92.9.1471

Brandehoff, R. (2025). Kanaka ʻŌiwi methodologies to honor Moʻōlelo, communities, Aloha ʻĀina, and make it Hānai. In M. L. Boucher, Jr. (Ed.), *Light a fire and watch it burn: Disruptive qualitative methods in education research* (pp. 1-9). Myers Education Press.

Campbell-Montalvo, R. A. (2020). Being QuantCritical of U.S. K-12 demographic data: using and reporting race/ethnicity in Florida Heartland Schools. *Race Ethnicity and Education, 23*(2), 180–199. https://doi.org/10.1080/13613324.2019.1679748

Campbell-Montalvo, R. (2021). Linguistic re-formation in Florida heartland schools: School erasures of indigenous Latino languages. *American Educational Research Journal*, *58*(1), 32-67. https://doi.org/10.3102/0002831220924353

Castillo, W. & Babb, N. (2024). Transforming the future of quantitative educational research: a systematic review of enacting quantCrit. *Race Ethnicity and Education, 27*(1), 1–21.

Castillo, W. & Strunk, K. K. (2025). *How to QuantCrit: Applying Critical Race Theory to quantitative data in education*. Routledge.

Catalyst California. (2024, August 15). *Why we must all be advocates when it comes to race data*. Retrieved February 2, 2026, from https://www.catalystcalifornia.org/blog/why-we-must-all-be-advocates-when-it-comes-to-race-data

Cepeda, R., & Combs, L. D. (2024) Disrupting monoracism, racism, and white supremacy: A systematic literature review on multiraciality in higher education. *Journal of Student Affairs*, *31*(2023-2024), 26–41.

Chang, M.J., Nguyen, M.H., & Chandler, K.L. (2015). Can data disaggregation resolve blind spots in policy making? Examining a case for Native Hawaiians. *AAPI Nexus: Policy, Practice, and Community*, *13*(1-2), 297-320. https://doi.org/10.17953/1545-0317.13.1.295

Charmaraman, L., Woo, M., Quach, A., & Erkut, S. (2014). How have researchers studied

multiracial populations? A content and methodological review of 20 years of research. *Cultural Diversity & Ethnic Minority Psychology, 20*(3), 336–352. https://doi.org/10.1037/a0035437

Chaudhari, P., & Pizzolato, J. E. (2008). Understanding the epistemology of ethnic identity development in multiethnic college students. *Journal of College Student Development, 49*(5), 443–458.. https://doi.org/10.1353/csd.0.0028

Cooper, J., Heron, T., & Heward, W. L. (2021). *Applied Behavior Analysis* (3rd ed.). Pearson.

Coupet, S. M. (2010). "Ain't I a parent?": Exclusion of kinship caregivers from debate over expansions of parenthood, *New York Review of Law and Social Change, 34*, 595-656.

Cruz, A. T., Palmer, C. A., Augustine, E. M., Casper, T. C., Dowshen, N., Elshholz, C. L., Mollen, C. J., Pickett, M. L., Schmidt, S. K., Stukus, K. S., Goyal, M. K. & Reed, J. L. (2024). Concordance of adolescent gender, race, and ethnicity: Self-report versus medical record data. *Pediatrics, 153*(2), 1-4. https://doi.org/10.1542/peds.2023-063161

el Zein, R. (2021). Introduction: Cultural constructions of race and racism in the Middle East and North Africa/Southwest Asia and North Africa. *Lateral (Island Lake), 10*(1). https://doi.org/10.25158/L10.1.11

Empowering Pacific Islander Communities (EPIC). (n.d.). *Our community.* EPIC. https://www.empoweredpi.org/our-community

Espinoza, K. J. C., Rincón, B. E., Drake, B. M., Harbin, J. J., & Ethelbah, K. K. (2025). Oceania in the desert: A QuantCrit analysis of the (under) counting of Native Hawaiian and Pacific Islander Students at an AANAPISI-HSI. *Journal of Research on Educational Effectiveness, 18*(2), 342-365.

Espiritu, Y. L. (1992). *Asian American panethnicity: Bridging institutions and identities*. Temple

University Press.

Fineman, M. (2006). Chapter 1. The meaning of marriage. In A. Bernstein (Ed.), *Marriage proposals: Questioning a legal status* (pp. 29-69). New York University Press.

Florida Legislature (2023). The 2023 Florida statutes (including special section C). https://www.leg.state.fl.us/Statutes/index.cfm?App_mode=Display_Statute&URL=0700-0799/0760/Sections/0760.80.html

Ford, K. S. (2019). Observer-identification: A potential threat to the validity of self-identified race and ethnicity. *Educational Researcher, 48,* 378–381. https://doi.org/10.3102/0013189X19860803

Garcia, N. M., López, N., & Vélez, V. N. (2018). QuantCrit: Rectifying quantitative methods through critical race theory. *Race Ethnicity and Education*, *21*(2), 149–157. https://doi.org/10.1080/13613324.2017.1377675

Gillborn, D., Warmington, P., & Demack, S. (2018). QuantCrit: Education, policy, 'Big Data' and principles for a critical race theory of statistics. *Race Ethnicity and Education, 21*(2), 158–179. https://doi.org/10.1080/13613324.2017.1377417

Ginorio, A. B., & Berry, P. C. (1972). Measuring Puerto Ricans' perceptions of racial characteristics. *Proceedings of the Annual Convention of the American Psychological Association, 7*(Pt. 1), 287–288.

Gogue, D. T. L., Poon, O. A., Maramba, D. C., & Kanagala, V. (2022). Inclusions and exclusions: Racial categorizations and panethnicities in higher education. *International Journal of Qualitative Studies in Education, 35*(1), 71-89. https://doi.org/10.1080/09518398.2021.1982045

Grosfoguel, R. (2004). Race and ethnicity or racialized ethnicities? Identities within global coloniality. *Ethnicities, 4*(3), 315-336. https://doi.org/10.1177/1468796804045237

Gutiérrez, R., & Dixon-Román, E. (2011). Beyond gap gazing: How can thinking about education comprehensively help us (re)envision mathematics education? In B. Atweh, M. Graven, W. Secada, & P. Valero (Eds.), *Mapping equity and quality in mathematics education* (pp. 21–34). Springer. https://doi.org/10.1007/978-90-481-9803-0_2

Hahn, R. A., Mulinare, J., & Teutsch, S.M. (1992). Inconsistencies in coding of race and ethnicity between birth and death in US infants. A new look at infant mortality, 1983 through 1985. *Journal of the American Medical Association, 267*(2), 259–63. https://doi.org/10.1097/00006254-199207000-00014

Harris, J. C. (2016). Toward a critical multiracial theory in education. *International Journal of Qualitative Studies in Education, 29*(6), 795–813. https://doi.org/10.1080/09518398.2016.1162870

Hsieh, B. & Kim, J. (2020) Challenging the invisibility of Asian Americans in education. *Review of Education, Pedagogy, and Cultural Studies, 42*(2), 95-103. https://doi.org10.1080/10714413.2020.176.0643

Iftikar, J. S. & Museus, S. D. (2018). On the utility of Asian critical (AsianCrit) theory in the field of education. *International Journal of Qualitative Studies in Education, 31*(10), 935-949. https://doi.org/10.1080/09518398.2018.1522008

Immigration and Nationality Act of 1965, Pub. L. No. 89-236, 79 Stat. 911 (1965).

Jones, N., Marks, R., Ramirez, R., & Ríos-Vargas, M. (2021, August 12). *Improved race, ethnicity measures reveal United States population much more multiracial*. U.S. Census Bureau. https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-

measures-reveal-united-states-population-much-more-multiracial.html

Junn, J., & Masuoka, N. (2008). Asian American identity: Shared racial status and political context. *Perspectives on Politics, 6*(4), 729–740. http://www.jstor.org/stable/20446825

Khanna, N. (2010) "If you're half Black, you're just Black": Reflected appraisals and the persistence of the one-drop rule, *The Sociological Quarterly, 51*(1), 96-121. https://doi.org/10.1111/j.1533-8525.2009.01162.x

Kina, L. & Dariotis, W. M. (2012). Miscegenating discourses: Critical contexts for mixed race Asian American art and identity. In L. Kina & W. M. Dariotis (Eds., pp. 3-21) *War baby/love child: Mixed race Asian American art*. University of Washington Press.

Lee, S. J. (1996). *Unraveling the model minority stereotype: Listening to Asian American youth.* Teachers College Press.

Lopez, M. H. (2013, October 22). Pew Research Center. *Three-fourths of Hispanics say their community needs a leader: Most Latinos cannot name one*. Pew Research Center. https://www.pewresearch.org/race-and-ethnicity/2013/10/22/3-hispanic-identity

Low, K. (2010, May 04). Federal Way schools refuse to guess student ethnicities. *Federal Way Mirror.* https://www.seattlepi.com/local/sound/article/Federal-Way-schools-refuse-to-guess-student-885395.php

Madrid, C. & Cepeda, J. (2019). Chamorro as the written official language in the Mariana Islands: The Garrido document of 1799. *Pacific Asia Inquiry, 10*(1), 117-140. https://www.uog.edu/schools-and-colleges/college-of-liberal-arts-and-social-sciences/pacific-asia-inquiry/V10

Malaney-Brown, V. K. (2020). On the path to multiracial consciousness: Reflections on my

scholar-practitioner journey in higher education. In M. P. Johnston-Guerrero & C. L. Wijeyesinghe (Eds.), *Multiracial experiences in higher education: Contesting knowledge, honoring voice, and innovating practice* (pp. 88–97). Stylus Publishing.

Massey, J. T. (1980). A comparison of interviewer observed race and respondent reported race in the National Health Interview Survey. Social Statistics Section. *Proceedings of the American Statistical Association*, 425-428.

Mills, K. A., & Unsworth, L. (2018). The multimodal construction of race: A review of critical race theory research. *Language and Education*, *32*(4), 313–332. https://doi.org/10.1080/09500782.2018.1434787

National Center for Education Statistics [NCES]. (n.d.). *Annual reports and information staff (Annual Reports): Allocations*. U.S. Department of Education Institute of Education Sciences. https://nces.ed.gov/surveys/annualreports/topical-studies/allocations

National Center for Education Statistics [NCES]. (2022). *Racial/Ethnic enrollment in public schools: Condition of Education* [Data set]. U.S. Department of Education, Institute of Education Sciences. Retrieved March 24, 2026, from https://nces.ed.gov/programs/coe/indicator/cge

National Center for Education Statistics [NCES]. (2011). NCES 2019-050. *Early childhood longitudinal study, kindergarten class of 2010-11 Fall Kindergarten Parent Interview*. In Data Collection Instruments. U.S. Department of Education, Institute of Education Sciences. Retrieved March 30, 2026, from https://nces.ed.gov/ecls/pdf/kindergarten2011/Fall_K_Parent_Interview.pdf

National Center for Education Statistics [NCES]. (2019). *User's manual for the ECLS-K:2011 kindergarten–fifth grade data file and electronic codebook*. U.S. Department of

Education, Institute of Education Sciences. Retrieved Feb 9, 2026, from https://nces.ed.gov/pubs2019/2019051.pdf

National Commission on Asian American and Pacific Islander Research in Education [CARE]. (2011). *The relevance of Asian Americans & Pacific Islanders in the college completion agenda.* Author. https://files.eric.ed.gov/fulltext/ED536943.pdf

Nguyen, B. M. D., Alcantar, C. M., Curammeng, E., Hernandez, E., Kim, V., Paredes, A., Nguyen, M., & Teranishi, R. T. (2017). *The racial heterogeneity project: Implications for higher education research, practice, and policy.* The Institute for Immigration, Globalization, & Education. https://files.eric.ed.gov/fulltext/ED583600.pdf

Office of Management and Budget. (2024). *Statistical Policy Directive No. 15: Standards for maintaining, collecting, and presenting federal data on race and ethnicity*. Federal Register. Retrieved January 10, 2026, from https://www.federalregister.gov/d/2024-06469

Okamoto, D. G. (2014). *Redefining race: Asian American panethnicity and shifting ethnic boundaries.* Russell Sage Foundation.

Omi, M., Nguyen, M. H., & Chan, J. (2020). Panethnicity and ethnic heterogeneity: The politics of lumping and disaggregating Asian Americans and Pacific Islanders in educational policy. In R. T. Teranishi, B. M. D. Nguyen, C. M. Alcantar, & E. R. Curammeng (Eds.), *Measuring race: Why disaggregating data matters for addressing educational inequality* (pp. 46–66). Teachers College Press.

Omi, M., & Winant, H. (2014). *Racial formation in the United States*. Taylor & Francis Group.

Parker, L. (1998). ‘Race is race ain't’: An exploration of the utility of critical race theory in qualitative research in education. *International Journal of Qualitative Studies in*

Education, *11*(1), 43–55. https://doi.org/10.1080/095183998236881

Parker, C. & Taychachaiwongse Teng, D. (2024, December 6-7). *Colonial obfuscation* [Paper presentation]. 14th International Conference on Education and Justice 2024, Honolulu, HI.

Poon, O. A., Dizon, J. P. M., & Squire, D. (2017). Count me in!: Ethnic data disaggregation advocacy, racial mattering, and lessons for racial justice coalitions. *Journal Committed to Social Change on Race and Ethnicity*, *3*(1), 91-124. https://doi.org/10.15763/issn.2642-2387.2017.3.1.91-124

Poon, O., Squire, D., Kodama, C., Byrd, A., Chan, J., Manzano, L., Furr, S., & Bishundat, D. (2016). A critical review of the model minority myth in selected literature on Asian Americans and Pacific Islanders in higher education. *Review of Educational Research, 86*(2), 469-502. https://doi.org/10.3102/0034654315612205

Rice v. Cayetano, 528 U.S. 495 (2000). https://supreme.justia.com/cases/federal/us/528/495/

Rivera, C. & Pennock-Roman, M. (1987). *Issues in race/ethnicity identification procedures in the National Assessment of Educational Progress, part I: A comparison of observer reports and self-identification.* Educational Testing Service. https://doi.org/10.1002/j.2330-8516.1987.tb00223.x

Ruiz, N. G., Noé-Bustamante, L., & Shah, S. (2023, May 8). Pew Research Center. *Diverse cultures and shared experiences shape Asian American identities.* Pew Research Center. https://www.pewresearch.org/race-and-ethnicity/2023/05/08/asian-identity-in-the-us/

Saenz, R., Hwang, S.-S., & Aguirre, B. E. (1994). In search of Asian war brides. *Demography*, *31*(3), 549–559. https://doi.org/10.2307/2061757

Schneider, Phillip A.A. (1980). Race/ethnic data collection in the Federal Civil Service. Survey

Research Methods Section, *Proceedings of the American Statistical Association*, 364-367. http://www.asasrms.org/Proceedings/papers/1980_075.pdf

Sojka, P. C., Maron, M. M., Dunsiger, S. I., Belgrave, C., Hunt, J. I., Brannan, E. H., & Wolff, J. C. (2025). Evaluation of reliability between race and ethnicity data obtained from self-report versus electronic health record. *Journal of Racial and Ethnic Health Disparities*, 12, 2200–2203. https://doi.org/10.1007/s40615-024-02041-w

Strunk, K. K. (2025). Against a critical quantitative orthodoxy: Proliferation of methodologies as a liberatory potential. *Inquiry: Critical Thinking Across the Disciplines, 34*(2), 520-537. https://doi.org/10.5840/inquiryct2025102334

Suzuki, B. H. (1977). Education and the socialization of Asian Americans: A revisionist analysis of the "model minority" thesis. *Amerasia Journal, 4*(2), 23–51. https://doi.org/10.17953/amer.4.2.x203l74863857108

Taitano, G. E. (n. d.). *Origin of Chamorro as an ethnic identifier.* Guampedia. https://www.guampedia.com/origin-of-chamorro-as-an-ethnic-identifier

U.S. Census Bureau. (2022). *2022 ACS 1-year public use microdata sample (PUMS)* [Data set]. 2022. https://nces.ed.gov/programs/digest/d23/tables/dt23_202.20.asp

U.S. Census Bureau. (n.d.). *Why we ask questions about race.* American Community Survey. https://www.census.gov/acs/www/about/why-we-ask-each-question/race

Vezaldenos, V. (2025). Quant(Multi)Crit: Situating the multiracial experience within quantitative critical race theory. *Inquiry: Critical Thinking Across the Disciplines, 34*(2), 538-568. https://doi.org/10.5840/inquiryct2025102335

Wanhalla, A. (2023). *Of love and war: Pacific brides of World War II*. University of Nebraska Press.

Wong-Campbell, J. P., & Ramrakhiani, S. H. (2025). Defining mixed-race college students: Multiracial (re) categorization and the visibility of graduation gaps. *Journal of Diversity in Higher Education, 18*(S1), S203-S212. https://doi.org/10.1037/dhe0000556

Yi, V., Mac, J., Na, V. S., Venturanza, R. J., Museus, S. D., Buenavista, T. L., & Pendakur, S. L. (2020). Toward an anti-imperialistic critical race analysis of the model minority myth. *Review of Educational Research, 90*(4), 542-579. https://doi.org/10.3102/0034654320933532

Zaslavsky, A. M., Ayanian, J. Z., & Zaborski, L. B. (2012). The validity of race and ethnicity in enrollment data for Medicare beneficiaries. *Health Services Research, 47*(3-2), 1300–1321. https://doi.org/10.1111/j.1475-6773.2012.01411.x

Zeiger, S. (2010). *Entangling alliances: Foreign war brides and American soldiers in the twentieth century*. New York University Press.

Zembylas, M. (2025). Rethinking positionality statements in research: from looking back to building solidarity. *International Journal of Research & Method in Education*, *48*(5), 466–478. https://doi.org/10.1080/1743727X.2025.2475762

Appendix A

Table A1

QuantCrit vs. Quant(Multi)Crit

QuantCrit Theory	Quant(Multi)Crit Theory
Centrality of Racism: Systemic racism operates both through formal and informal mechanisms, whether overtly or covertly. To disrupt the insidiousness of racism, researchers must critically interrogate the processes and practices associated with quantitative data.	**Centrality of Racism, Monoracism, and Colorism**: Scholars should consider how colorism might be accounted quantitatively, paying close attention to whether and how their research may reproduce monoracism through racial classification decisions.
Numbers are not neutral: Quantitative data is steeped in ideologies of white supremacy and positivism; however, numbers are neither any more or less objective compared to qualitative data (Castillo & Strunk, 2025; Vezaldenos, 2025).	**Numbers perpetuate a monoracist racial categorization system:** Scholars should consider the ways their research decisions might contribute to multiracial identities being erased or misrepresented.
Categories are neither neutral nor natural: Race is a social construction with categories used to maintain hierarchies. This requires deeper examinations of how categories can perpetuate inequities, while proposing alternatives for dealing with data representing marginalized communities in more just ways.	**Critique (Multi)racial categories: for 'race,' read '(mono)racism:** Researchers should examine subtle ways monoracism might be influencing how racial data is constructed and used in education research while refusing to conflate all Multiracial experiences as being the same.
Voice and insight: data cannot speak for itself: Research is influenced by a variety of internal and external factors (e.g., researcher positionality and funding). Thus, it is imperative to challenge dominant narratives by centering marginalized voices in your work.	**Center Multiracial voice and insight:** Challenging the monoracial status quo requires centering Multiracial voices and experiences.
Using numbers for social justice: Numbers should be used toward social justice aims as quantitative research findings are influenced by scholars' methodological decisions.	**Using numbers for social justice:** To disrupt the monoracist status quo, researchers must engage a social justice orientation toward quantitative data.

Note. Table adapted from Vezaldenos (2025). QuantCrit tenets and definitions displayed here are from Gillborn et. al (2018, p. 169) and Castillo & Babb (2022, p. 3); Quant(Multi)Crit tenets and definitions are from Vezaldenos (2025, p. 11).

Appendix B

ECLS-K Parent Interview Protocol

Figure B1

Hispanic/Latino Ethnicity Item (ECLS-K Item FSQ.190)

QUESTION TEXT:

{Are you/is {NAME} Hispanic or Latino?}

FIRST NAME	HISPANIC/LATINO
{Display HH Member Name}	[ ]
{Display HH Member Name}	[ ]
{Display HH Member Name}	[ ]

HELP TEXT:

Hispanic or Latino: A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race.

CODES:

1 YES
2 NO
REFUSED
DON'T KNOW

PROGRAMMER INSTRUCTIONS:

Bold "Hispanic or Latino" in help text.

CAPI ROSTER INSTRUCTIONS: Display in column 1 each person enumerated on the household roster (FSQ.020) Who is the focal child, respondent, mother figure (Code ‘1’ at FSQ.130 or code '3' at FSQ.180), or father figure (code ‘2’ at FSQ.130 or code '4' at FSQ.180).

If no mother or father figures in the household (no household members with a code ‘1’ at FSQ.130, code '3' at FSQ.180, code '2' at FSQ.130, or code '4' at FSQ.180), Display in column 1 the focal child, the respondent, and the respondent’s spouse/partner (household member selected at FSQ.120, if any).

NOTE: If the respondent is a mother or father figure, only display his/her name once.

CAPI MATRIX INSTRUCTIONS:

1. Display first name, last name, and age of the person, using information from the current round household matrix.

2. The cursor should be positioned on the first blank field.
3. The first column of the matrix (first name) is read only (see roster instructions above).
4. Display “are you” if the subject of the question is the respondent. else, display “is {name}” using the name of the household member that is the subject of the question.
5. Display '{Are you/is {NAME}} Hispanic or Latino?" and "1 = YES 2 = NO" when cursor is positioned in the hispanic column of the matrix.
6. Refused and don't know are allowed for Hispanic/Latino.
7. Cursor will move from Hispanic/Latino column to race column in FSQ.195 for same person and then will move to Hispanic/Latino column in FSQ.190 for next person, etc. The cursor will move in this fashion until all fields are completed.
8. Interviewer cannot leave the matrix until all answer fields are accounted for.

Note. Parents/caregivers are first asked whether their child is “Hispanic or Latino.”

Figure B2

Race Item (ECLS-K Item FSQ.195)

QUESTION TEXT:

What is {your/[NAME]}'s race? You may name one or more races to indicate what {you/[NAME]} consider(s) yourself/himself/herself to be.

{IF "HISPANIC" OR "LATINO" PROBE: Is that White Hispanic, Black Hispanic, both, or something else?}

{IF RESPONDENT CONTINUES TO SAY "HISPANIC" OR "LATINO" AFTER USING THE PROBE ABOVE, CODE AS "DON'T KNOW."}

FIRST NAME	RACE
{Display HH Member Name}	[ ]
{Display HH Member Name}	[ ]
{Display HH Member Name}	[ ]

HELP TEXT:

American Indian or Alaska Native: A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

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.

Black or African American: A person having origins in any of the Black racial groups of Africa.

Native Hawaiian or Other Pacific Islander: A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

White[a]: A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

CODES:

1 AMERICAN INDIAN OR ALASKA NATIVE
2 ASIAN
3 BLACK OR AFRICAN AMERICAN
4 NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER
5 WHITE

PROGRAMMER INSTRUCTIONS:

Bold "American Indian or Alaska Native," "Asian," "Black or African American," "Native Hawaiian or Other Pacific Islander," and "White" in first instance only of each in help text.

Display “your,” “you,” “consider,” and “yourself” if the subject of the questions is the respondent. else, display the name of the subject of the questions, “considers” and “himself” if the subject of the questions is male. else, display the name of the subject of the questions, “considers” and “herself” if the subject of the questions is female.

CAPI ROSTER INSTRUCTIONS: Display in column 1 each person enumerated on the household roster (FSQ.020) Who is the focal child, respondent, mother figure (code ‘1’ at FSQ.130 or code '3' at FSQ.180), or father figure (code ‘2’ at FSQ.130 or code '4' at FSQ.180).

If no mother or father figures in the household (no household members with a code ‘1’ at FSQ.130, code '3' at FSQ.180, code '2' at FSQ.130, or code '4' at FSQ.180), Display in column 1 the focal child, the respondent, and the respondent’s spouse/partner (household member selected at FSQ.120, if any).

NOTE: If the respondent is a mother or father figure, only display his/her name once.

CAPI MATRIX INSTRUCTIONS:

1. Display first name, last name, and age of the person, using information from the current round household matrix.
2. The cursor should be positioned on the first blank field.
3. The first column of the matrix (first name) is read only (see roster instructions above).
4. Display "What is {your/{NAME}'s} race?," “If...something else,” “If...don’t know,” "code all that apply," and answer categories when cursor is positioned in the race column of the matrix.
5. Refused and don't know are allowed for race.
6. Cursor will move from the race column (FSQ.195) back to the Hispanic/Latino column (FSQ.190) for the next person in the matrix, if there is another household member to be asked about. the cursor will move in this fashion until all fields are completed.
7. Interviewer cannot leave the matrix until all answer fields are accounted for.

Note. Parents/caregivers are asked to select their child’s race from five race categories, with the option of selecting multiple categories.

[a] The ECLS-K uses racial categories established by OMB guidelines (NCES, 2011), which included persons from the Middle East or North Africa in the white race category. In 2024, a new “Middle Eastern or North African” (MENA) ethnicity category was developed in addition to the newly mandated use of a combined race and ethnicity question format (OMB, 2024). While these updates represent incremental progress, they are still structurally deficient in their exclusion of racially and ethnically diverse groups within the broad MENA label, such as Black Arabs and Armenian Americans (Arab American Institute, 2024). Moreover, the term “Middle East” has been debated as “recent activist energy has turned to encourage the use of a less colonized alternative, SWANA—Southwest Asia and North Africa” (el Zein, 2021, Notes section).