

Racial Category Usage in Education Research: Examining the Publications from AERA Journals

Dominique J. Baker 

University of Delaware

Karly S. Ford 

Pennsylvania State University

Samantha Viano 

George Mason University

Marc P. Johnston-Guerrero 

University of Denver

How scholars name different racial groups has powerful salience for understanding what researchers study. We explored how education researchers used racial terminology in recently published high-profile peer-reviewed studies. Our sample included 1,427 original empirical studies published in the nonreview AERA journals from 2009 to 2019. We found that two thirds of articles used at least one racial category term, with an increase from about half to almost three quarters of published studies between 2009 and 2019. Other trends include the increasing popularity of the term Black, the emergence of gender-expansive terms such as Latinx, the popularity of the term Hispanic in quantitative studies, and the paucity of studies with terms connoting missing race data or including terms describing Indigenous and multiracial peoples.

Keywords: *ethnicity, race, sociology of knowledge, descriptive analysis*

Discussions referencing specific racial groups are ubiquitous in education research.¹ These descriptions are often superficial, such as when a quantitative study includes race indicators as covariates, or tangential, as is the case where race is mentioned but not critically explored (Garcia & Mayorga, 2018; James, 2001; Ladson-Billings, 2012; Tabron & Thomas, 2023). However, the ways in which we name and describe different racial groups have powerful salience for understanding what researchers believe and what they study (Gillborn et al., 2018; Ladson-Billings, 2012; Ma et al., 2007; Turner et al., 2024).

Racialization, the application of racial meaning to a “relationship, social practice, or group” in the United States is a complex process that, while socially constructed, can result in tangible consequences (Omi & Winant, 2014, p. 111). Race is a sociopolitical and historical categorization structure that can differ across cultures and over time (e.g., Omi & Winant, 2014; Roth et al., 2023; Viano & Baker, 2020). As noted in Howell and Emerson (2017), “no one can

identify ‘true’ racial classifications, because they do not exist” (p. 15). These shifts in racial classification systems should not be considered natural or value neutral. The variations in sets of racial categories have often sprung from the need to maintain White supremacy, leading, within the United States, to enslavement, the stealing of land and resources from Indigenous peoples, Jim Crow segregation, explicit barring of Asian people from immigration and citizenship, and eugenics, among other policy decisions (Omi & Winant, 2014; Saperstein & Penner, 2012). Incorporating racial group status into research without deeper engagement with the reasoning behind the categories thus can reinforce harms and may lead to research that is used to oppress groups of people. For instance, as described by O’Connor, Lewis, and Mueller (2007), a variable indicating Black racial identification is often included with a long list of covariates in statistical models with little theorization or engagement with prior literature to explain why the researchers think it useful to include the information. In doing so, identifying or being



perceived as Black is reduced to a category whose membership is mediated through what are seen as sociocultural deficiencies while also obfuscating the process of racialization that sorts groups of people into the category of being Black (Ladson-Billings, 2012; O'Connor et al., 2007; Omi & Winant, 2014; Zuberi & Bonilla-Silva, 2008). Consequently, the decisions about how to name and describe race in education research have powerful salience for interpretation of the findings.

Even though racial data are expected measures in most analytic frameworks (e.g., “race without racism”; Harper, 2012, some areas of the education research community have little critical engagement with understanding contemporary usages of these terms in our research studies (Denton & Deane, 2010; Johnston-Guerrero, 2017). Better understanding of how education researchers use language to describe racial groups is paramount to broader efforts to create an inclusive education research community (Galvez & Muñoz, 2020; Salinas, 2020) as well as communicating education research more effectively.

Although uncommon in education, researchers in other fields (e.g., biomedicine and demography) regularly examine how racial data are collected, categorized, and used (e.g., Caulfield et al., 2009; Rachul et al., 2011; Shanawani et al., 2006; Stevens et al., 2015). For example, Lee (2009) reviewed National Cancer Institute–supported publications for their terminology related to race, finding that these studies commonly invoked racial terminology but rarely described definitions or operationalization. The education research community similarly could benefit from an explicit examination of how race is being categorized/discussed in scholarship because this, in turn, shapes how future researchers employ racial categorization and frame the implications of their findings.

Unlike prior work that critically analyzes how education research often avoids recognizing racism as the cause of racial gaps (Harper, 2012; Kohli et al., 2017), this kind of exploration of the literature seeks to understand what specific language education researchers are using to describe racial categories. How researchers write about racial categories is likely heterogeneous within fields for a variety of reasons, including the expectations of journal editors, common practices of reporting by method, and trends in how others are writing about racial categories. First, both Harper (2012) and Kohli et al. (2017) recommend that journals play a role in reinforcing requirements or norms around the discussion of racial categories, supporting the hypothesis that the journal publishing the article might partially determine how racial categories are described (Lee, 2009; Ma et al., 2007). Second, racial categorizations are dynamic, with categories changing based on shifting social understandings about the boundaries between groups and fluctuating expectations around what is considered appropriate terminology (Denton & Deane, 2010; Ma et al., 2007; Williams, 1999). For

example, the term *Asian American* was created in the 1960s as a political resistance tactic (e.g., Ishizuka, 2018), Irish people were not originally considered White on immigration to the United States (e.g., Brodtkin, 1998), and, this year, the U.S. Office of Management and Budget (OMB) announced Middle Eastern and North African as a new racial category (previously included under White) for the standard racial categories for the federal government (Wang, 2024). One might expect to see some shifts in language that might, nevertheless, fail to perpetuate fieldwide.

Third, the differences in how qualitative and quantitative researchers theorize and analyze racial differences are well documented, with substantial criticism of the ways in which quantitative research often reinforces racial hierarchies instead of supporting justice-focused efforts to eliminate racial inequality (Kohli et al., 2017; O'Connor et al., 2007; Turner et al., 2024; Zuberi, 2001). These critiques particularly highlight the erasure of student groups with smaller population sizes, such as Indigenous students (Shotton et al., 2012). To our knowledge, these observations have not yet been paired with an analysis of how qualitative or quantitative studies operationalize racial categories. Exploring racial category usage across all three of these complicating factors would allow for a heightened understanding of how educational research has recently conceptualized these categories and potential reasons why these categorizations vary.

In this study we explore how educational researchers have used racial terminology in published peer-reviewed studies between 2009 and 2019. We focus on original research published in journals from the American Educational Research Association (AERA, <https://www.aera.net/Publications/Journals>) because these journals attract a wide variety of articles of different orientations and publish articles at the forefront of education research. Our work addresses the following research questions:

1. To what extent are there trends over time in the rate of inclusion of racial categories in published education research overall, within journal, and by research method?
2. What terminology does published education research use to describe racial categories? Does that terminology differ by journal, research method, or over time?

By systematically exploring these published articles, we seek to make the following contributions: (1) quantify how often researchers acknowledge or omit race in education research, (2) provide a baseline understanding of which racial categories education researchers use and their frequency, (3) investigate differences among journals, (4) explore trends over time for how language on racial groupings is changing, and (5) understand differences across methodologies. In doing so, we seek to spark a conversation about how we, as a community of education researchers, can

represent racial groups in our research in ways that are both representative of our participants and reflective of broader efforts toward antiracism and decolonization.

Racial Categorization/Classification

Classifying and categorizing are part of the human social experience (Bowker & Star, 2000). Humans draw lines and classify most everything in our social world—we divide land into towns, counties, states, and nations in similar fashion to how we categorize people by language, ancestry, religion, beliefs, and other characteristics. Categories are social phenomena, laden with the contextual, political, and social understandings of the people who create and use them. Consequently, categories play a central role in our lives. They are often treated as static when, in fact, categorization is quite dynamic. Race categories in the United States, and around the world, are impacted by histories of settler colonialism, enslavement, xenophobia, and systemic anti-Black, anti-Indigenous, and anti-immigrant legacies (Mills, 1997; Omi & Winant, 2014; Saucier & Woods, 2018; Zuberi & Bonilla-Silva, 2008). These significant social forces shape how race categories are constructed and have real, material impacts on peoples' lives—their access to healthcare, schooling, employment, and housing (Smedley & Smedley, 2005).

Increasingly, race scholars in sociology and political science are recognizing race as having multiple dimensions (Morning, 2009; Rockquemore et al., 2009; Roth, 2018; Sen & Wasow, 2016). A person has a racial identity—a rich tapestry of being and experiencing. A person has one or more racial categories with which they self-identify when confronted with a survey that prescribes discrete options (Rockquemore et al., 2009; Johnston et al., 2014). A person is racially appraised (or identified by observers) when they enter a classroom, walk down a street, or are seen at a hospital (López et al., 2018; Telles, 2014). A person's racial identity, racial category, and “street race” may all align or may differ significantly depending on the context (López et al., 2018). In many cases, there is fluidity among and between these dimensions of race over the life course and depending on the context.

In education scholarship, researchers employ racial categories as shorthand to describe a rich, complex web of racializing experiences. Although these categories help to highlight group-level processes (places where systems of oppression are faced), the categories themselves hold power. Education researchers who use race categories without reflection risk reifying notions of innate racial difference. For instance, when studying the so-called Black–White test score gap, it can be easy for categorization to suggest that the gap is produced by differences innate to Black and White students rather than differences produced via systemic racism in and outside the schoolhouse (Ladson-Billings, 2006,

2007; O'Connor et al., 2007; Welner & Carter, 2013). Aligning with this concern, scholars have found evidence that using the framing of achievement gaps causes an increase in the negative perception of Black students among members of the general public and K–12 teachers (Quinn, 2020; Quinn & Desruisseaux, 2022). Researchers too often measure the effects of racism and attribute them to racial difference—reifying race and racism (Sewell, 2016; Zuberi, 2001). In summary, racial categories are created by people, either imposed or claimed, and are related to political and material interests. Racial categories have tangible impacts on peoples' lives, including the kinds of resources and opportunities that are available to them.

Other Fields' Consideration of Racial Categorization

Although uncommon in education, systematic and critical reviews of which racial categories are being used by researchers is common in other fields. Demography, epidemiology, media studies, political science, and criminology have regularly examined how racial data are collected, categorized, and used in research (e.g., Covington, 1995; Garcia, 2017; Gomez & Glaser, 2006; Hahn et al., 1996; Kanakamedala & Haga, 2012; Kelley et al., 1996; Megyesi et al., 2011; Shrikant & Sambaraju, 2021). We will describe two examples from other fields to illustrate this point.

From epidemiology, Ma et al. (2007) systematically reviewed every research article published in the *Annals of Internal Medicine*, the *Journal of the American Medical Association*, *Lancet*, and the *New England Journal of Medicine* between 1999 and 2003 ($n = 1,152$) for the included racial categories. They found that researchers referred to White using 16 terms, 13 terms for Black, 16 terms for Asian, and 11 terms related to Hispanic ancestry (Ma et al., 2007). After describing their results by journal and year, Ma et al. connected their systematic review of racial categories to larger issues of policy and practice in epidemiology. They wrote that to meet the goals set forth by the National Institute of Health (2005) and the American College of Physicians (Groman & Ginsburg, 2004), reflection and consistency across studies about how racial categories are used were a necessary “initial step to closing the health care gap” (Ma et al., 2005, p. 577).

A second illustrative example comes from demography. Stevens et al. (2015) analyzed the historical censuses of the United States, Canada, and Australia to examine how the racial categories have shifted over time. These three nations were selected because they all share a similar history of having Indigenous inhabitants, settler colonialism, and huge waves of global immigration in the 20th century. The authors identified three trends shared across these nations. First, new categories and new groupings of categories are added over time and reflect immigration patterns. For example, in the United States, the category of “Hispanic” was introduced in

1980, Canada began to use the concept of “visible minority” in 1990, and Australia began collecting parents’ birthplace as well as birthplace of the individual. Second, European ethnic groups (e.g., Italian and German) were once used as different racial categories. As racial difference among White people of direct European descent became less socially salient, a single White category became more common (although in Canada and Australia, the national signifiers “Canadian” and “Australian” are increasingly popular).

Third, each nation moved through several permutations of how to classify and categorize people with mixed-race heritage. In the first century of each census, the enumerator (door-to-door census taker) would visually observe, assess, and record a single race of each person. Later, government agencies provided decision rules for how to classify the race of the individual. In the United States, the race of the father was used in 1970 for a mixed-race person, but then the race of the mother was deemed a better indicator in 1980. Prior to the 1970s in Australia, the government asked fractions to be provided (e.g., “one half Aboriginal, one half Chinese”; Stevens et al., 2015, p. 25). Between 1990 and 2000, the census categories for the three countries changed to accommodate the selection of two or more races. Stevens et al. (2015) concluded that these three countries have different conceptions of race measurement, but they all share common confusion because of the volatility and complexity of racial identification.

Scholars in fields outside of education value these previous analyses of how race categories vary because where race boundaries are drawn shapes how people understand their social realities and also shapes the lessons scholars can take away from the research. We use these insights to inform our analysis of published education research, as described in the following section.

This Study

This review of prior research has argued that it would be shortsighted for education researchers to take racial categories for granted. Doing so makes invisible power and political/social mores, leads to conceptual and methodologic confusion, and, in some cases, can lead to faulty conclusions that can misinform policy/practice. This review of prior work reiterates that categories reflect the political and social interests of people and that, historically, race categories have been used to turn difference into a way to exclude and justify the oppression of people whose origins are not White/European.

In addition to drawing from theories and prior work on categorization/classification, we ground our study in the sociology of knowledge that focuses on studying the daily, taken for granted, shared assumptions and rules about social life (e.g., Garfinkel, 1967). Sociologists of knowledge argue that the deepest insights about the social world could be

observed in studying what others may view as mundane, such as the social norms, routines, and beliefs guiding daily life (Berger & Luckmann, 1967; Boutros, 2024; Garfinkel, 1967; Go, 2020; Kuhn, 1970; Merton, 1972; Morning, 2011; Riley et al, 2021; Roberts, 2011; Swidler & Ardit, 1994; Zuberi & Bonilla-Silva, 2008). Scholars can examine people’s lived experiences and motivations that gird daily life or the behaviors that result from those views. We focus on the latter in the topic motivating this study—the race categories researchers routinely use. As Omi and Winant (2014) highlight, “race operates as a ‘common-sense’ concept, a basic component of social cognition, identity, and socialization. . . . Race seems obvious and in some ways superficial” (p. 4). Therefore, the sociology of knowledge offers a valuable opportunity to investigate the racial classification system used by education scholars.

In line with this theory tradition, we are denaturalizing the categories by examining them closely and carefully not as objective reality but as decisions that people make and remake until they shape our social reality. What categories are being used in education research? By what kinds of methodologic researchers? How has this changed over time? These questions may seem basic, though, given our theoretical frame, troubling the basic is where the deep social insights reside. Indeed, our research questions are foundational because the social process of categorization precedes all other research processes (Hirschman et al., 2016).

Given that race categories are not “natural” but rather the products of social and political decisions, we argue that researchers must pay attention to which categories are used and how they change, approaching these categories as political and social creations rather than “facts of biology and/or fate” (Gillborn et al., 2018, p. 172). Of note, when examining classification through this lens, it is also imperative to consider who has the power and authority to determine which categories are appropriate in a given time and place because this can play a central role in the creation of categories and their impact (e.g., Collins, 2015; Morning, 2009, 2011; Roberts, 2011; Swidler & Ardit, 1994). Because there is no “true” set of racial categories for any society, we believe that the role played by key stakeholders (e.g., the OMB, that determines which racial categories will be used by the federal government, and editors of academic journals) must be part of how we make meaning of the use of racial categories in this study. In contemporary education research, most empirical analyses use fixed categories for race without much consideration or reflection. Although education journals have not engaged in this kind of examination of which race categories are used, we demonstrate that this need not be the case as other fields carefully examine how scholars use racial categories in research. This study aims to bring this kind of systematic analysis and careful thinking about racial categories to education research.

Research Methods

We provide an overview of our data-collection and analysis methods here. We include a positionality statement in supplementary material for this article (see Supplement A in the online version of the journal).

Data Collection

A member of the research team compiled a list of all potential research articles published from 2009 to 2019 in the AERA journals that do not exclusively publish reviews: *AERA Open*, *American Educational Research Journal (AERJ)*, *Educational Evaluation and Policy Analysis (EEPA)*, *Educational Researcher (ER)*, and *Journal of Educational and Behavioral Statistics (JEBS)*.² We include original empirical research to allow the authors of those articles more control over the racial categorization, since review articles likely would use the language of the articles they were reviewing. This search resulted in 1,623 articles.³

A different research team member created a coding frame focused on (1) whether the article included original empirical research, (2) key facts about the article (e.g., keywords, first author's academic affiliation), (3) research methodology, (4) racial categories included for Asian/Pacific Islander, Black, Native American, Latinx, White, two or more races, and missing/unknown, and (5) whether the article studied U.S.-based populations (see Supplement B in the online version of the journal for the full coding frame, which included additional items not analyzed in this study). For the racial categories, we began with a list of categories based on the U.S. Census during the analytical timeframe (2009–2019). As an example, for the Asian/Pacific Islander racial category, the form prompted: "Race/ethnicity category(ies) for Asian and Pacific Islander (used anywhere in the paper, check all that apply): (1) Asian, (2) Asian American, (3) Native Hawaiian, (4) Pacific Islander, (5) N/A, (6) Other (Free response)." Coders selected all options that applied to each article with original empirical research.

To begin, all authors coded five randomly selected articles and met to reach consensus on all items and revise the coding frame based on the initial coding.⁴ Once the coding frame was finalized, the four authors split the remaining articles and coded them separately. Two authors completed their coding with the aid of trained research assistants. Once the research team completed coding the articles, a research team member created a 10% random sample from the list of articles assigned to each of the four authors (132 articles). A trained research assistant who had not conducted any of the original coding then double-coded all the articles in the random sample, and we compared the two sets of coding to assess interrater reliability. The coding had an overall reliability of 93% (the primary and secondary coders had identical responses for 93% of their codes) with reliability by coder ranging from 84% to 93%. Because the 84% reliability

was an outlier (the next lowest was 91%), two members of the research team completed a second round of coding for all the articles originally coded by the member with the lowest reliability rating. The first author then reconciled the original and second rounds of coding (either by retaining codes that matched across both coders or reviewing the article herself and making a final decision).

Once we removed articles that did not include original empirical research, our final dataset included 1,427 articles. Most articles (1,267; 88%) analyzed U.S. domestic data. The number of articles published per year has been increasing from 85 in 2009 to 221 in 2019 (includes the addition of *AERA Open* in 2015). The overwhelming majority of empirical articles used some type of quantitative method (~80%). Finally, *AERJ* published the most articles during the analytical time period (456), although *AERA Open* has published an extraordinary number since it came into existence in 2015, with 224 articles (*ER* has 200, *JEBS* 262, and *EEPA* 285).

Analysis Methods

With the final analytic set of articles coded, we created measures of the different racial categories. Creating these categories was an iterative process that involved the first author creating measures based on categories in the coding frame as well as the text written as free response by individual coders. The first author would then present these categories, along with the full list of all raw codes, to the research team to discuss refinement of the analytic categories. Based on this iterative process, completed three different times, we created a final set of analytic categories that we outline in Table 1. Beyond the categories based on the U.S. Census, we included options such as "non-Hispanic" (e.g., "non-Hispanic Black students"), gender-expansive terms for Latinx (e.g., "Latino/a" and "Latino/a/x"), and ethnic group, nationality, or region for Asian or Pacific Islanders (e.g., "Vietnamese," "Filipino," and "Chinese").⁵ We created binary variables that equal 1 if the article mentions a term within each category.⁶

We descriptively investigated the analytic set of articles using summary statistics focused on how frequently articles used terms for different racial categories.⁷ In addition to the racial categories, we included indicators of the journal outlet (*AERA Open*, *AERJ*, *EEPA*, *ER*, or *JEBS*), year published, and research method (i.e., qualitative, quantitative, or mixed). We then explored how the frequency of different racial categorizations differed by journal, year published, and method. We conducted all the analysis focused on specific racial categories solely on the articles that at least partially used data from the United States. We added this additional sample restriction for the analysis due to the variation in the sociopolitical context for race and ethnicity across countries (Marquardt & Herrera, 2015; Stevens et al., 2015).

TABLE 1
Analytical Categories

No.	Asian or Pacific Islander	Black	Indigenous	Latinx	Missing	Two or more	White
1	Any Asian or Pacific Islander category is used	Any Black category is used	Any Indigenous category is used	Any Latinx category is used	Any race missing category is used	Any two or more races category is used	Any White category is used
2	"Asian" is used	"Black" is used	"Native American" is used	"Hispanic" is used	"Unknown" is used	"Two or more races" is used	"White" is used
3	"Asian American" is used	"African American" is used	"American Indian" is used	"Latino" is used	"Missing" is used	"Mixed race" is used	"Caucasian" is used
4	"Native Hawaiian" or "Hawaiian" is used	"Non-Hispanic" is used	"Alaska Native" or "Alaskan Native" is used	"Latinx" is used	Decline to report is used	"Multiracial" is used	"Non-Hispanic" is used
5	"Pacific Islander" is used	—	"Indigenous" is used	Gender-expansive term is used	—	"Multiethnic" is used	"Eur" is used (for European or Eurasian)
6	"Non-Hispanic" is used	—	"Non-Hispanic" is used	—	—	—	—
7	Ethnic group, nationality or region is used	—	—	—	—	—	—

Note. For region or country within the Asian or Pacific Islander category, we included the following options: Vietnamese, Filipino, Chinese, Japanese, Korean, Malay, Indian, East, South, Hmong, Aleutian, Lao, Maori, Nepalese, Bengali, Pakistani, Sri Lankan, Cambodian, Guamanian, Thai, Indonesian, Fijian, Marshallese, Polynesian, Tahitian, Samoan, and Desi. For gender-expansive terms within the Latinx category, we included the following options: Latino/a, Latino(a), Latinas(os), Latino/a/x, Latinx, Latina/o, Latinas/Latinos, Latinas/as, and Latinos/Latinas. For decline-to-report terms within the missing category, we included the following options: decline, respond (e.g., did not respond or not responding), prefer (e.g., prefer not to say), indicate (e.g., not indicated), not report (e.g., not reported), and not provided. We determined the options based on the coded responses for the articles included in the sample.

Findings

In this section we address both research questions while examining the use of any racial categories and then specific terms.

Any Racial Category

We first review the trends in published education research using any racial category. Approximately two thirds of all the empirical articles published from 2009 through 2019 in AERA journals included at least one racial category. Figure 1 shows that a larger share of more recent articles used any racial category, from 53% in 2009 to 73% in 2019. Still, there was significant variation when we examined article characteristics and use of racial categories. A little more than 80% of the articles published by *AERJ* and *EEPA* used any racial category, with *AERA Open* following closely (75%). These journals published noticeably more articles that used any racial category compared with *ER* (63%) and *JEBS* (13%).

Turning to temporal trends by outlet, Figure 2 is a heat plot that uses a color gradient to visually show the percentage of articles that used at least one racial category over time by journal outlet. The 0% to 10% range is shown in yellow, with the color gradient transitioning to green tones after 10%, followed by a transition to blue and then purple tones at 50% and above. The darkest purple tone indicates that 100% of articles in that journal in that timeframe included racial categories. In *AERA Open*, which began publishing in 2015, consistently more than 50% of its articles used a racial category. This was similar for both *AERJ* and *EEPA*. We can visually observe the increase in the proportion of *ER* articles using racial categories over time with the cells between 2009 and 2011 being more of a yellow-greenish hue and the articles between 2016 and 2019 having a purple hue. In contrast, *JEBS* decreased its share of articles that use racial categories, with cells between 2009 and 2013 having a greenish hue compared with articles after 2014 generally being yellow. In 2009, 18% of *JEBS* articles used any racial category, and this decreased to 8% by 2019. Although all research methodologies had a majority of articles that used a racial category, quantitative articles had a smaller share than qualitative or mixed-methods articles (64% vs 75% and 71%). There is little evidence of a temporal trend across the different methods.

Specific Racial Categories

Now that we have provided an overview of any racial category use, we turn to the specific categories we created for different racial groups looking solely at domestic data. Figure 3 shows the share of articles in each year that mentioned any of the terms in each racial category (as described

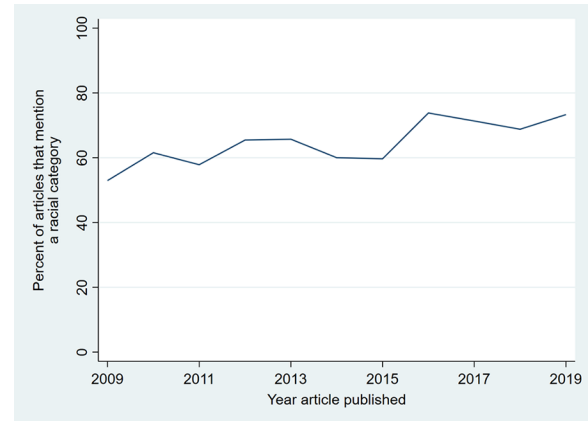


FIGURE 1. *Percentage of Articles that Used at Least One Racial Category Over Time*

Note. Overall annual percentage of empirical articles that mentioned any racial category by publication year.

in Table 1). Overall, use of the different racial categories has increased over time. The largest share of articles mentioned terms for White, Black, and Latinx categories and then the Asian or Pacific Islander category, with Indigenous and two or more races categories following. Terms denoting race missing were the smallest percentage every year.

Turning to the individual categories, we expanded on the categories listed in Table 1 by examining data on the percentage of articles using the terms in each row within columns (e.g., the proportion of articles using each term for Asian or Pacific Islander) overall, by year, by journal, and by method in Tables 2 through 8 with one table for each column of Table 1. We structure our discussion of the results by racial category alphabetically using our selected term representing each broader category. Table 2 shows the share of articles that mentioned any terms in the Asian or Pacific Islander category. Each row presents percentages for each subsample, starting with the overall total, and then each publication year, the journal outlets, and research method. Interpreting one estimate as an example, we found that ~43% of articles used any Asian or Pacific Islander term, with the majority term being “Asian” (42%) as compared with “Asian American” (8%), “Native Hawaiian” (4%), “Pacific Islander” (11%), “non-Hispanic” (e.g., “non-Hispanic Asian,” <1%), and ethnic group/nationality/region (3%). As shown in Figure 3, the use of any Asian or Pacific Islander term increased over time from 35% of articles including any Asian/Pacific Islander term in 2009 to 50% of articles in 2019. Although *AERJ* had the largest share of articles using any terms for the Asian category (58%), *AERA Open* had the largest share of articles using “Native Hawaiian” and “Pacific Islander” (8% and 18%, respectively). Quantitative research had the largest share of articles using “Asian,” “Native Hawaiian,” “Pacific Islander,” and

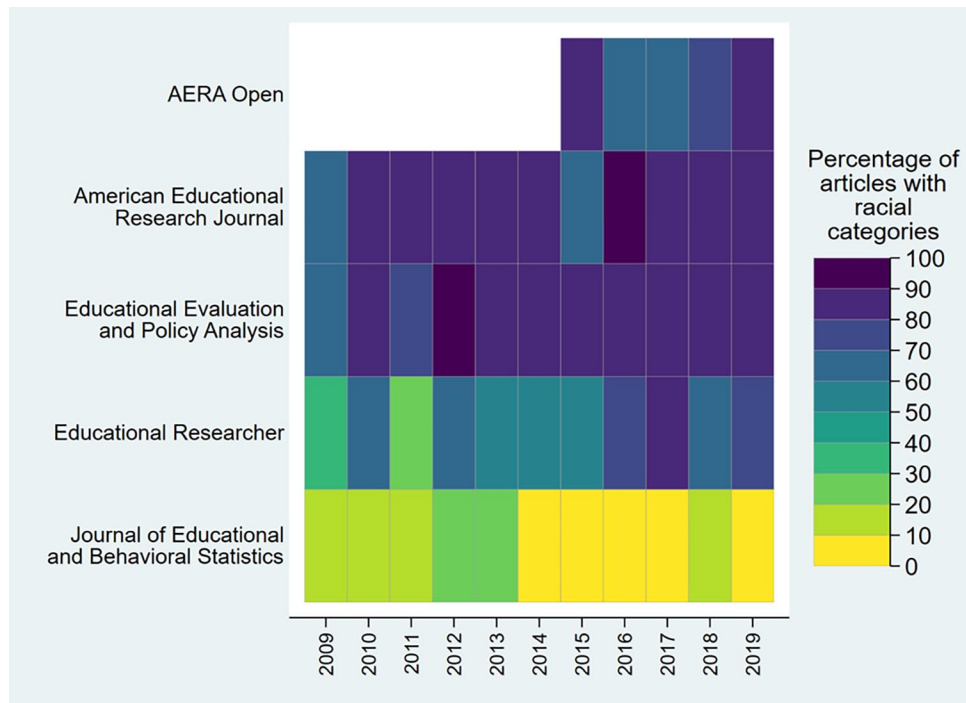


FIGURE 2. *Percentage of Articles that Used at Least One Racial Category Over Time by Journal*

Note. Rows are separated by journal. Columns are separated by publication year. Cell color gradient visually represents the percentage of articles within the respective journal and year that mentioned any racial category. *AERA Open* published its first article in 2015.

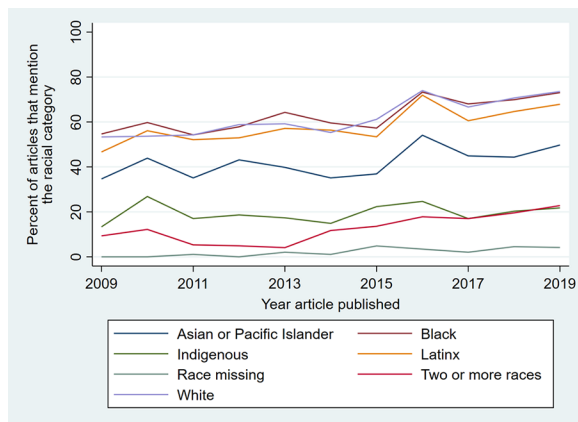


FIGURE 3. *Percentage of Articles that Mentioned at Least One Term in Each Racial Category*

Note. Overall annual percentage of empirical articles that mentioned at least one term in each respective racial category (listed in alphabetical order) by publication year.

“non-Hispanic,” whereas qualitative research was more likely to use the terms “Asian American” and ethnic group/nationality/region-specific terminology.

Table 3 shows the Black categories. We found that ~65% of all articles used any term for Black, with a general increase in use over time from 55% in 2009 to 73% in 2019. The largest share of articles used the term “Black” compared with “African American” and “non-Hispanic” (51% vs 36% and

2%, respectively). We can see that the increase in the mention of any category for Black was driven by the term “Black,” which appeared in only 36% of articles in 2009 and rose to 63% in 2019, whereas use of “African American” declined from a high of 43% of articles in 2010 to a low of 29% in 2018. Turning to journal outlets, we found that *AERJ* generally had the highest percentage of articles that mentioned any term for the Black category. Although articles published in *AERA Open* and *EEPA* were more likely to use the term “Black,” articles published in *AERJ* were more likely to use the term “African American.” While quantitative research had the smallest percentage of articles using any Black category term (64% compared with 69% of qualitative and 68% of mixed methods), those articles had the highest percentage for “non-Hispanic” (2%). Articles using quantitative research methods also had the largest share of articles that used the term “Black” (nearly 52% vs almost 48% for qualitative and 46% for mixed methods).

We explore the Indigenous category in Table 4. About a fifth of publications used any term for Indigenous, with the most popular term being “Native American” (11%), closely followed by “American Indian” (10%), with “Alaskan Native” (5%), “Indigenous” (0.3%), and “non-Hispanic” (0.2%) being less popular. Unlike the use of terms for Black and Asian, we did not observe a consistent increase in terms for Indigenous, with the peak percentage of articles including an Indigenous category occurring in 2010 (26%) and fluctuating between 15% and 25% thereafter. There were similar inconsistencies

TABLE 2

Term Use for the Asian or Pacific Islander Category

Factor	Any Asian or Pacific Islander category	Asian	Asian American	Native Hawaiian	Pacific Islander	Non- Hispanic	Ethnic group/ nationality/region
Time period							
Overall	43.33	42.38	8.13	3.55	11.05	0.47	2.92
2009	34.67	34.67	6.67	0	4	1.33	2.67
2010	43.9	43.9	12.2	1.22	15.85	0	1.22
2011	35.11	34.04	8.51	1.06	7.45	0	3.19
2012	43.14	43.14	5.88	0.98	9.8	0	5.88
2013	39.8	38.78	10.2	3.06	8.16	2.04	4.08
2014	35.11	34.04	5.32	1.06	6.38	0	2.13
2015	36.89	34.95	4.85	2.91	9.71	0	2.91
2016	54.11	52.74	11.64	3.42	12.33	0.68	5.48
2017	44.9	43.54	7.48	6.12	9.52	0	2.04
2018	44.36	42.86	7.52	4.51	15.79	0	2.26
2019	49.74	49.22	8.29	7.77	15.54	1.04	1.04
Journal							
<i>AERA Open</i>	54.45	54.45	8.9	7.85	18.32	1.57	0.52
<i>AERJ</i>	58.15	55.64	14.79	4.26	13.53	0.75	6.52
<i>EEPA</i>	47.39	46.64	4.48	1.12	8.96	0	1.87
<i>ER</i>	39.67	39.67	7.07	5.43	12.5	0	2.72
<i>JEBS</i>	5.78	5.78	0.89	0	1.78	0	0
Method							
Qualitative	44.62	41.54	12.82	2.56	9.23	0	7.69
Quantitative	43.5	42.91	7.28	3.84	11.71	0.59	2.17
Mixed	35.71	35.71	7.14	1.79	5.36	0	0

Note. Each column reports the percentage of articles that used the respective column's term(s) for each row's subsample. "Mixed" under "Method" refers to mixed-methods research.

TABLE 3

Term Use for the Black Category

Factor	Any Black category	Black	African American	Non- Hispanic
Time period				
Overall	64.64	50.91	36.46	2.05
2009	54.67	36	40	1.33
2010	59.76	47.56	42.68	2.44
2011	54.26	38.3	36.17	0
2012	57.84	43.14	33.33	1.96
2013	64.29	50	39.8	2.04
2014	59.57	44.68	31.91	1.06
2015	57.28	43.69	32.04	1.94
2016	73.29	56.16	39.73	3.42
2017	68.03	53.74	37.41	0
2018	69.92	60.9	28.57	4.51
2019	73.06	62.69	39.38	2.59
Journal				
<i>AERA Open</i>	77.49	64.4	39.27	5.76
<i>AERJ</i>	80.7	58.4	56.64	1.75

(continued)

TABLE 3 (CONTINUED)

Factor	Any Black category	Black	African American	Non- Hispanic
<i>EEPA</i>	77.99	63.81	37.69	0.37
<i>ER</i>	59.78	51.63	26.63	2.17
<i>JEBS</i>	13.33	10.22	4.89	1.33
Method				
Qualitative	69.23	47.69	60.51	0.51
Quantitative	63.58	51.77	31.4	2.36
Mixed	67.86	46.43	44.64	1.79

Note. Each column reports the percentage of articles that used the respective column's term(s) for each row's subsample. "Mixed" under "Method" refers to mixed-methods research.

by term with the use of "Native American" and "American Indian" fluctuating up and down by year, although the use of "Alaskan Native" and "Indigenous" increased over time. Similar to the other racial categories, although *AERJ* had the largest share of articles using any term, *AERA Open* had the largest percentages for "American Indian" and "Alaskan Native." Quantitative research methods again had the highest

TABLE 4
Term Use for the Indigenous Category

Factor	Any Indigenous category	Native American	American Indian	Alaskan Native	Indigenous	Non-Hispanic
Time period						
Overall	19.81	10.81	10.26	5.13	0.32	0.16
2009	13.33	9.33	5.33	2.67	0	0
2010	26.83	15.85	15.85	6.1	0	0
2011	17.02	11.7	5.32	0	0	0
2012	18.63	10.78	9.8	5.88	0	0
2013	17.35	11.22	8.16	3.06	0	1.02
2014	14.89	9.57	5.32	1.06	0	0
2015	22.33	12.62	9.71	3.88	0	0
2016	24.66	11.64	12.33	6.85	1.37	0
2017	17.01	9.52	8.84	6.8	0	0
2018	20.3	7.52	13.53	7.52	0.75	0
2019	21.76	10.88	13.47	7.25	0.52	0.52
Journal						
<i>AERA Open</i>	22.51	8.9	15.18	9.42	0.52	0.52
<i>AERJ</i>	26.57	16.79	12.78	5.76	0.75	0.25
<i>EEPA</i>	20.52	12.31	7.46	2.99	0	0
<i>ER</i>	20.11	8.15	13.59	7.61	0	0
<i>JEBS</i>	4.44	2.22	2.22	0.89	0	0
Method						
Qualitative	20.51	14.36	7.69	3.59	2.05	0
Quantitative	19.78	10.14	10.93	5.61	0	0.2
Mixed	17.86	10.71	7.14	1.79	0	0

Note. Each column reports the percentage of articles that used the respective column’s term(s) for each row’s subsample. “Mixed” under “Method” refers to mixed-methods research.

percentage for “non-Hispanic” as well as “American Indian” and “Alaskan Native,” and only qualitative articles used the term “Indigenous.”

Turning to the Latinx category, Table 5 presents the trends. Although 60% of articles used any term for Latinx, the highest percentage of articles used “Hispanic” followed by “Latino” (49% and 23%, respectively). While use of each term generally (although inconsistently) increased over time, “Latinx” and the gender-expansive category had an exceptional increase in representation in 2019 (from 3% in 2018 to 17% in 2019 for Latinx and 4% in 2018 to 20% in 2019 for any gender-expansive term). Unlike the prior racial categories, *AERA Open* published the largest share of articles mentioning any Latinx term. Although articles published in *AERJ* typically mentioned the previous categories at the highest rates, in the Latinx category, only “Latino” was the highest for *AERJ*. Continuing the reverse, quantitative methods had the highest share of articles using any term for Latinx categories. This reversal in the trend was overwhelmingly driven by use of the term “Hispanic,” whereas qualitative and mixed-methods articles were more likely to use the other terms.

Table 6 includes the usage of terms for the race-missing category. The most consistent trend was that terms for the

TABLE 5
Term Use for the Latinx Category

Factor	Any Latinx category	Hispanic	Latino	Latinx	Gender expansive
Time period					
Overall	59.91	49.49	22.65	3.39	6.08
2009	46.67	37.33	16	0	2.67
2010	56.1	52.44	23.17	0	2.44
2011	52.13	43.62	24.47	0	3.19
2012	52.94	45.1	17.65	0	3.92
2013	57.14	48.98	20.41	0	4.08
2014	56.38	41.49	22.34	0	4.26
2015	53.4	39.81	21.36	0	2.91
2016	71.92	60.96	25.34	2.05	5.48
2017	60.54	52.38	23.81	2.04	2.72
2018	64.66	55.64	21.8	3.01	3.76
2019	67.88	52.33	26.42	17.1	19.69
Journal					
<i>AERA Open</i>	76.44	69.11	21.99	8.38	10.47
<i>AERJ</i>	74.19	53.63	40.85	5.51	9.77

(continued)

TABLE 5 (CONTINUED)

Factor	Any Latinx category	Hispanic	Latino	Latinx	Gender expansive
<i>EEPA</i>	72.39	65.3	18.28	1.12	1.87
<i>ER</i>	53.8	45.11	16.85	1.09	7.07
<i>JEBS</i>	10.67	10.22	0.89	0	0
Method					
Qualitative	56.92	37.44	40.51	8.72	12.31
Quantitative	60.63	52.26	18.8	2.36	4.72
Mixed	57.14	41.07	30.36	3.57	8.93

Note. Each column reports the percentage of articles that used the respective column's term(s) for each row's subsample. "Mixed" under "Method" refers to mixed-methods research.

TABLE 6

Term Use for the Race-Missing Category

Factor	Any race-missing category	Race unknown	Race missing	Decline/no response
Time period				
Overall	2.45	1.1	0.63	0.95
2009	0	0	0	0
2010	0	0	0	0
2011	1.06	1.06	0	0
2012	0	0	0	0
2013	2.04	1.02	1.02	0
2014	1.06	1.06	0	0
2015	4.85	2.91	0.97	0.97
2016	3.42	0.68	0.68	2.05
2017	2.04	0.68	0.68	0.68
2018	4.51	2.26	1.5	0.75
2019	4.15	1.55	1.04	3.11
Journal				
<i>AERA Open</i>	4.19	1.57	1.05	1.57
<i>AERJ</i>	2.51	0.75	0.75	1.5
<i>EEPA</i>	2.61	1.49	0.75	0.37
<i>ER</i>	2.72	2.17	0	1.09
<i>JEBS</i>	0.44	0	0.44	0
Method				
Qualitative	1.54	1.54	0.51	0
Quantitative	2.46	0.98	0.69	0.98
Mixed	5.36	1.79	0	3.57

Note. Each column reports the percentage of articles that used the respective column's term(s) for each row's subsample. "Mixed" under "Method" refers to mixed-methods research.

race-missing category were infrequently used, appearing in only 2% of articles. The share of articles mentioning a race-missing category increased over time from 0% to 4%–5% (depending on the year). Articles published in *AERA Open* were the most likely to include a race-missing category, although the highest percentage for any individual term is

still less than 2%. We observed few differences between qualitative and quantitative articles, with mixed-methods articles being the most likely to include a race-missing category.

We present the trends for the two-or-more-races category in Table 7. Although 14% of articles included any two-or-more-races category, the most popular term was "multiracial" at 9%. Unlike most of the other categories, use of any term for the two-or-more-races category consistently increased from 2014 (whereas the other categories generally increased but did have periods of decrease). Similar to the Latinx category, *AERA Open* articles had the largest share of terms for the two-or-more-races category. Qualitative articles were most likely to mention any two-or-more-races category, although qualitative and quantitative articles had similar proportions using the term "multiracial."

The results for the final category, White, are presented in Table 8. Overall, 64% of articles used any White term, with the majority term being "White" (61%) as compared with "Caucasian" (6%), "non-Hispanic" (5%), and "European" (2%). As shown in Figure 3, the use of any White term increased over time, with *AERJ* and *AERA Open* publishing the largest shares of articles with any White term (81% and 77%, respectively). This increase was concentrated among the terms "White" and "non-Hispanic" with no consistent changes in the use of "Caucasian" and "European." Similar to the overall use of any racial term, qualitative and mixed-methods research used White terms more frequently than quantitative research. Similar to some of the other racial categories, although most of the White category terms were used less frequently in quantitative research, "non-Hispanic" was used more frequently in quantitative and mixed-methods research than in qualitative research (5% for both compared with 2%).

Discussion

In this study we conducted an analysis of how education researchers used racial terminology in published peer-reviewed studies appearing in five AERA journals between 2009 and 2019. Based on our findings (described earlier), we now highlight the complexities and overall contributions of the study. We organize the discussion around trends over time, less common terminology, how our findings relate to the U.S. Census, and differences across journals. Throughout the discussion, we rely on our understanding of the construction of racial categorization to explain how our observations on researchers' use of racial terminology could be driven by the larger national sociopolitical context.

This discussion of results has several inherent limitations to consider when interpreting our findings. First, we can only observe what authors wrote in the final publication. In other words, we do not observe their data or their original manuscript prior to the revision process. We do not know

TABLE 7

Term Use for the Two-or-More-Races Category

Factor	Any two-or-more-races category	Two or more races	Mixed race	Multiracial	Multiethnic
Time period					
Overall	13.97	1.97	1.58	8.68	1.26
2009	9.33	0	2.67	4	0
2010	12.2	1.22	1.22	6.1	2.44
2011	5.32	0	1.06	5.32	1.06
2012	4.9	0	0	4.9	0
2013	4.08	0	0	4.08	0
2014	11.7	0	2.13	6.38	1.06
2015	13.59	0.97	1.94	9.71	1.94
2016	17.81	1.37	2.74	9.59	1.37
2017	17.01	2.72	0	10.88	0
2018	19.55	5.26	1.5	12.78	1.5
2019	22.8	5.18	3.11	12.95	3.11
Journal					
<i>AERA Open</i>	26.7	5.24	2.62	15.18	2.62
<i>AERJ</i>	20.05	2.51	3.26	12.03	2.26
<i>EEPA</i>	8.58	0.75	0.37	6.72	0
<i>ER</i>	11.41	1.63	0.54	7.61	1.09
<i>JEBS</i>	0.89	0	0	0.44	0
Method					
Qualitative	18.46	3.59	3.08	8.21	2.56
Quantitative	13.19	1.67	1.28	8.96	1.08
Mixed	12.5	1.79	1.79	5.36	0

Note. Each column reports the percentage of articles that used the respective column's term(s) for each row's subsample. "Mixed" under "Method" refers to mixed-methods research.

TABLE 8

Term Use for the White Category

Factor	Any White category	White	Caucasian	Non-Hispanic	European
Time period					
Overall	63.93	60.77	5.76	4.5	1.66
2009	53.33	52	5.33	1.33	0
2010	53.66	52.44	4.88	4.88	2.44
2011	54.26	46.81	8.51	2.13	1.06
2012	58.82	54.9	6.86	3.92	1.96
2013	59.18	54.08	7.14	4.08	2.04
2014	55.32	52.13	3.19	2.13	3.19
2015	61.17	59.22	2.91	0.97	1.94
2016	73.97	70.55	6.16	9.59	4.11
2017	66.67	64.63	4.08	2.72	0.68
2018	70.68	66.92	8.27	7.52	0
2019	73.58	71.5	5.7	5.7	1.04
Journal					
<i>AERA Open</i>	77.49	73.3	7.85	8.38	1.57
<i>AERJ</i>	81.45	75.44	10.03	5.76	4.26
<i>EEPA</i>	73.13	72.39	2.99	1.49	0

(continued)

TABLE 8 (CONTINUED)

Factor	Any White category	White	Caucasian	Non-Hispanic	European
<i>ER</i>	61.96	59.78	4.89	4.89	0
<i>JEBS</i>	12	11.11	0.44	2.22	0.44
Method					
Qualitative	69.74	65.64	7.18	2.05	5.13
Quantitative	62.7	59.65	5.31	4.92	0.98
Mixed	66.07	64.29	8.93	5.36	1.79

Note. Each column reports the percentage of articles that used the respective column's term(s) for each row's subsample. "Mixed" under "Method" refers to mixed-methods research.

whether the language we observe is a result of decisions during data collection that might not have been within the control of the authors. It is also possible that racial category language is influenced by the revision process such that the authors might have preferred different terminology than what was published. Second, our findings could be partially a reflection of journal word limits. *AERJ* accepting the longest manuscripts could mean more space for including discussion of racial categories even if they are not a core aspect of the analysis. In journals such as *ER* that accept shorter articles, these terms might only appear in supplementary material (which we did not analyze). We did not specifically measure the density or quality of discussion or use of racial category terminology, so the higher rate of racial category usage in *AERJ* could reflect room for longer tables that include covariates. Other journals could plausibly include a higher density of articles that authentically engage in discussions on racial categorization, which would not have been captured in our coding framework.

Trends in Categorical Usage Over Time

There is much to learn from the trends in usage of particular racial categories. We found an overall increase over time in term usage across all categories. All categories seemed to match a similar upward trend, with fluctuations over time. However, the two-or-more-races category uniquely showed a consistent upward trajectory after 2013. Given the fact that it was not until 2010–2011 that the U.S. Department of Education mandated that institutional data collection meet OMB Directive 15 guidelines for allowing students to report two or more races (Renn, 2009), perhaps this distinct trend occurred because of the potential lag in getting these data (or even data from the 2010 U.S. Census) before then moving toward publication. Moreover, the trend likely also reflects a steady increase in the representation of and consciousness around multiracial people across education (Harris, 2016; Howard, 2018), even with continued debate about the utility of such grouping for civil rights laws (Hernández, 2018). One additional insight is that "multiracial" was the most used term within two or more races,

which proves interesting given that naming practices are widely contested for this group. For example, some multiracial people identify as mixed, biracial, or, more specifically, "Blasian" (i.e., Black and Asian) or "Mexipina" (i.e., Mexican and Filipina).

Our findings also documented a trend in various categories used to describe Black populations, with a steady increase in the specific term "Black" while "African American" decreased over the analytic time period. This trend aligns with increasing consciousness around racial injustice and solidarity among Black peoples across the diaspora, especially given that recent immigrants from the African continent and Caribbean, for instance, might see themselves as Black but not African American (Fries-Britt et al., 2014).

Similarly, the gender-expansive term "Latinx" first appeared in AERA journals in 2016, where only 2.05% of articles included it. By 2019, "Latinx" skyrocketed to 17.1% appearance across all journals except for *JEBS*, along with a substantial increase in the usage of other gender-neutral terms such as "Latina/o." Our findings did not suggest that these gender-expansive terms are necessarily replacing other terms representing Hispanic/Latino groups because "Hispanic" also increased and "Latino" held steady across the 10 years.

Relatedly, we found a general upward trend in the use of any White category, from 53.3% in 2009 to 73.6% in 2019. Although this matches the general increasing trend in all categories, we highlight the importance of naming White/Whiteness as a racial category instead of deeming it the default or norm (Sue, 2004).

Lower-Frequency Terminology

We highlight the significantly smaller number of articles in our dataset that engaged with the Asian or Pacific Islander, Indigenous, two-or-more-races, and race-missing categories. This lack could be reflective of demographic representation, a function of the prominence of the Black–White binary, or researchers' notions of the necessary sample size for sufficient statistical power in quantitative analysis (even

though the necessary sample size to uphold statistical assumptions is actually fairly low compared with the samples in the quantitative research we reviewed). The model minority myth that Asian Americans (and Pacific Islanders as a result of collapsing groups together) are widely successful (Jang, 2018) may contribute to a larger narrative that research is not necessarily needed to support this community (compared with others). Moreover, Indigenous population sizes often have relegated them to being merely an asterisk in research studies noting that their small sample size precludes them from being separated out in the analysis (Shotton et al., 2012), further emphasizing the colonial erasure of Indigenous peoples in education research.

Our findings showed that the percentage of articles engaging the race-missing category is *incredibly* small. This low engagement could be because, within the K–12 education administrative data landscape, there are no missing-race or unknown students (Ford, 2019). Ford (2019) explained how school administrators must assign a racial category (through observer identification) to students if they do not provide a racial self-identification. Yet, within the higher education sector, race-unknown students are often clustered at the most and least selective institutions (Ford et al., 2021, 2022). More research is needed to understand the individual-level motivations and incentives for opting out as well as the organizational practices of collecting/reporting data. For instance, Renn (2004) found a pattern of multiracial identity termed “extraracial,” where students were opting out of racial categorization in an attempt to deconstruct race and exist beyond racial categories. However, other research suggests that the race-unknown category for college students is largely White students (e.g., Ford & Holland, 2020), which may signify a desire to distance oneself from Whiteness or lack of knowledge that “White” is a racial identity.

Although the term is not extremely popular, appearing in about 5% of published educational research, continual use of the term “Caucasian” is troubling given the research highlighting the term’s problematic history rooted in White supremacy. Mukhopadhyay (2018) traced the origins of this term to 18th-century Europeans desiring to classify peoples in an emerging “racial science,” with Johann Blumenbach popularizing the term for Europeans to have origins in the Caucasus mountains because he saw the light-skinned people of this region as the most beautiful and ideal type of humans (in “God’s image”). Blumenbach attributed value and character to these groupings, which emboldened the racial hierarchy with “Caucasians” on top and all others denigrated.

In the United States, key legal battles around which peoples could hold citizenship eventually led to “Caucasian” not just being a sociopolitical category but also a legal one with much consequence. The early 1920s U.S. Supreme Court cases *Ozawa v. United States* and *United States v. Thind* demonstrate the variability of justifications used to police

and solidify the boundaries of Whiteness. While Japanese-born Takao Ozawa was denied naturalization because, despite his white skin color, his race was not deemed “Caucasian,” the court later ruled against Bhagat Singh Thind that despite his “Caucasian” or Aryan origins, his brown skin meant that he was not White (Haney López, 1997). Though the term “Caucasian” has been popularized as a polite or scientific term (Saini, 2019), it is likely that people who continue to conduct research using this term have little to no idea of its racist history. However, much more consciousness has been raised about the term, and calls have been made to discontinue its usage in alignment with a broadening body of critical Whiteness studies (Matias & Boucher, 2021) that spotlights power dynamics associated with Whiteness. More attention is needed among education researchers to expose this term’s racist pseudoscience origins to spur discontinuation of its use.

Following Census Categories

We found that “non-Hispanic” was used more often as a qualifier when describing White categories (4.5%) and, to a lesser extent, Black categories (2.1%). Although this practice seems like an accurate reflection of the specific issues related to the two-part Hispanic ethnicity versus race question used by the U.S. Census, it also, in some ways, mimics the policing of the boundaries of Whiteness and who counts as White, especially because we observed less use of the preface “non-Hispanic” with Black populations despite warranting further nuance for Afro-Latinx people (Dache et al., 2019). Moreover, the miniscule usage of “non-Hispanic” with various Asian or Pacific Islander or Indigenous categories needs more attention. This dichotomy could lead to erasure of those who are both Hispanic and Asian, Pacific Islander, or Indigenous. For instance, many Latin American countries have had long histories of immigration from Asian countries (e.g., Japanese in Peru and Filipinos in Mexico; Hu-DeHart & López, 2008). How can our current understandings of racial categories better capture this population? Moreover, how do the current racial categories reveal a more extensive history as well as the long-standing effects of colonization in education?

The finding that “non-Hispanic” is used primarily by quantitative and mixed-methods researchers across all racial groups is also important to note. This is most likely because of alignment with how the U.S. Census collects its survey data given the separated Hispanic ethnicity versus race questions. For qualitative research, this modifier of “non-Hispanic” may not be needed (and looks to not be used as often), illuminating the intricacies of racial category terms and the variability in usage methodologically. For instance, “Hispanic” is used more in quantitative (52.22%) than qualitative (37.44%) studies, whereas “Latinx” is used more in qualitative (8.72%) than quantitative (2.37%) studies. As

mentioned earlier, the specific term “Latinx” is becoming more widely used across education and in popular culture (Salinas, 2020), though there has also been some recent pushback. Given that the origins of the term “Latinx” are in Latin America, people in the United States may not as easily recognize or accept it. For instance, Pew researchers found that although 23% of surveyed Hispanics had heard of the term, only 3% used it to describe themselves (Noe-Bustamente et al., 2020). Moreover, there have been critiques that “Latinx” is unpronounceable in Spanish or is an elite term used by academics, despite its origins being from community activists and many users recognizing the “x” as relating more to their Indigenous roots (Salinas, 2020). Our study showed the large increase in both “Latinx” and gender-expansive terms in 2019, demonstrating that researchers likely desire to recognize more expansive notions of gender but do not find “Hispanic” to be a good alternative (Viano & Baker, 2020). It will be interesting to see how this trend continues or whether “Latinx” might be replaced with Salinas’ (2020) recommendation of “Latin*” or “Latine” (an option reflecting Spanish grammar conventions; Slemp, 2020) as more inclusive and disruptive alternatives, especially as identities become further negotiated.

Overall, the wide variability in usage of different categories suggests that education researchers are not actually strictly following the U.S. Census categories. Instead, it is more likely that they are sticking closely to the categories used on various surveys, some of which align with the Census while others may not. Yet researchers have the ability to change the categories they use to better align with the lived experiences of minoritized people. For example, the Census does not use “Caucasian,” yet this term continues to be used in education research.

Journal Differences: Why Is JEBS an Outlier?

Our study found differences across journals in usage of different racial terms despite all being education research journals within the same professional association (AERA). *AERA Open* published the largest share of articles mentioning any Latinx term and specifically “Latinx.” *AERA Open* articles also had the largest share of terms for the two-or-more-races category as well as the specific terms “American Indian,” “Alaska Native,” “Native Hawaiian,” “Pacific Islander,” and “multiracial.” Given the historically fast review process of *AERA Open*, it may allow unique opportunities to more quickly adapt to changes in terminology. *AERJ* generally had the most racial term use except for the Latinx, two-or-more-races, and race-missing groups. Future studies could explore potential reasons for these differences, such as editorial board demographics or the editor statements that are published in the journals.

By not naming racial categories, education research tends to take a color-evasive approach (Annamma et al., 2017). This perspective might explain why *JEBS* had low rates of

using racial categories. However, this approach perpetuates the false idea that the field of statistics is objective and neutral (Gillborn et al., 2018; Zuberi, 2001) when, in fact, statistics was born out of the 19th-century eugenics movement (Saini, 2019; Zuberi, 2001; Zuberi & Bonilla-Silva, 2008). Social science research has a troubling history of using statistics to reinforce racial hierarchy, with many core tools of educational psychology and statistics (e.g., IQ tests), created by eugenicists, still in use today (Zuberi & Bonilla-Silva, 2008). *JEBS* is uniquely positioned to help guide the field in ways to incorporate race and racism into statistical models. As long as educational and behavioral statistics ignore this problematic history by maintaining a false stance of neutrality, so too will continue the commodification and, in this case, erasure of racial categories.

Conclusion

This study unveils the usage of racial categorization in high-profile published education research and, in doing so, uncovers several key findings related to changes in language usage, differences across journal/methodology, and growth areas for future published research. To the extent that the lack of discussion of racial category terminology in education research has led to problems such as the continued use of “Caucasian” or the low rates of using terms related to the two-or-more-races, Indigenous, and missing-race categories, this study opens the door for more careful reflection on what is included and excluded in our empirical writing. Future research on racial categorization in published education research could delve deeper into the density of discussion on racial categorization and the research interests/affiliations of writers, include a wider set of journals, or use content analysis to explore the reasoning given for the naming and inclusion of certain racial groups. We are encouraged by many of the trends we observed in this study related to higher usage of racial categories. Still, without future research, it remains unclear whether shifts in language use are happening concurrent with shifts in deep engagement with the constructs of race and racism. Understanding how racialization occurs within published education research hopefully can lead to better fieldwide norms on how to approach conducting and publishing research, which might then create better research evidence that can be used more easily by practitioners and policy actors.

It is difficult for us to provide concrete recommendations for the entire field of education based on this study, which is part of a broader research project examining the use of racial categories in education research (e.g., Viano et al., 2024). Certain recommendations we could provide might be more useful for qualitative research versus quantitative research, and adding even more nuance, there are also differences within methodologic traditions (e.g., quantitative research that uses secondary data versus data that are collected by the research team). Therefore, our primary recommendation,

beyond what we share in Viano et al. (2024), is that scholars who are interested in including race or racism as part of their research should ensure that they have educated themselves on the process of racialization (we have included several references throughout this paper that are useful starting points for this exploration). It would be useful for scholars to explore the rich empirical and theoretical research on race and racism within the societies they study. Of note, critical scholars have been doing deeply thoughtful research incorporating race and racism for a significant period of time and have provided overviews of research trends and ways to incorporate deeper introspection into the research process, often based on methodological design. At a bare minimum, scholars should have a reason for the inclusion of racial categories in their research as well as a reason for the set of categories used in their work (beyond “this is how the data arrived from the data provider”).

We know that questioning the use of various categorical terms does not necessarily connect with the material realities of racial disparities in education. Categories provide opportunities for pride, community building, solidarity, and coalitions across minoritized racial groups. However, we simultaneously acknowledge how these racial categories are rooted in White supremacy and an oppressive legacy, thus perpetuating systems of power in education research. Researchers must apply stewardship in research design decision points regarding identity category and term usage.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Open Practices Statement

The data and analysis files for this article can be found at <https://doi.org/10.3886/E209334V1>.

ORCID iDs

Dominique J. Baker  <https://orcid.org/0000-0001-8570-9681>
 Karly S. Ford  <https://orcid.org/0000-0002-2781-7761>
 Samantha Viano  <https://orcid.org/0000-0002-9229-3597>
 Marc P. Johnston-Guerrero  <https://orcid.org/0000-0001-9551-5186>

Notes

1. We use “racial” to include the main U.S. Census categories for race as well as the category “Hispanic, Latino, or Spanish origin.” We recognize that this final category is nominally an ethnicity in the United States, but with changes to the U.S. Census planned to fully incorporate this category as a racial group (Wang, 2024)

and the reality that this “ethnicity” is actually treated functionally as a product of a racialization process (Hitlin et al., 2007; Omi & Winant, 2014; Viano & Baker, 2020), we deem it appropriate to treat the “Hispanic” category as a racial category. For this reason, we generally use the term “racial” throughout this paper to refer to racialized groupings as experienced within the U.S. context.

2. Each AERA journal serves a distinct purpose and audience focused on education research. *AERA Open* publishes open-access articles, *AERJ* publishes full-length articles, *EEPA* publishes full-length policy-focused articles, *ER* publishes short articles, and *JEBS* publishes statistics and methods articles. For more information, go to <https://www.aera.net/Publications/Journals>. Note that AERA publishes two journals focused on reviews, *Review of Educational Research* and *Review of Research in Education*, which we exclude due to our restriction on articles featuring primary data analyses.

3. We exclude any introductory articles for special issues or new editorial teams, reviews/essays, policy forums or briefs, errata, software reviews, book reviews, and tutorials.

4. As an example of revision to the coding frame, we added additional options for the categories for Black individuals based on the articles we coded.

5. For gender-expansive terms within the Latinx category, we included the following options: Latino/a, Latino(a), Latinas(os), Latino/a/x, Latinx, Latina/o, Latinas/Latinos, Latinos/as, and Latinos/Latinas. For region or country within the Asian or Pacific Islander category, we included the following options: Vietnamese, Filipino, Chinese, Japanese, Korean, Malay, Indian, East, South, Hmong, Aleutian, Lao, Maori, Nepalese, Bengali, Pakistani, Sri Lankan, Cambodian, Guamanian, Thai, Indonesian, Fijian, Marshallese, Polynesian, Tahitian, Samoan, and Desi.

6. While the U.S. Census through 2020 considers Middle Eastern and North African (MENA) people as part of the White category, scholars and communities disagree (e.g., Maghbouleh et al., 2022). Because of this, we did not count the mentions of MENA categories in our set of articles within any of the analytic racial categorization (this affected five articles).

7. We do not include inferential statistics given that the data we study are a census of all education articles published in the nonreview AERA journals.

References

- Annamma, S. A., Jackson, D. D., & Morrison, D. (2017). Conceptualizing color-evasiveness: Using dis/ability critical race theory to expand a color-blind racial ideology in education and society. *Race Ethnicity and Education*, 20(2), 147–162. <https://doi.org/10.1080/13613324.2016.1248837>
- Berger, P., & Luckmann, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge*. Doubleday.
- Boutros, M. (2024). Antiracism without races: How activists produce knowledge about race and policing in France. *Social Problems*, 71(1), 1–17. <https://doi.org/10.1093/socpro/spac011>
- Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT Press.
- Brodin, K. (1998). *How Jews became white folks and what that says about race in America*. Rutgers University Press.
- Castillo, W., & Gillborn, D. (2022). *How to “QuantCrit:” practices and questions for education data researchers and users*

- (EdWorkingPaper 22-546). Annenberg Institute at Brown University. <https://www.edworkingpapers.com/ai22-546>
- Caulfield, T., Fullerton, S. M., Ali-Khan, S. E., Arbour, L., Burchard, E. G., Cooper, R. S., & Daar, A. S. (2009). Race and ancestry in biomedical research: Exploring the challenges. *Genome Medicine, 1*(1), 1–8. <https://doi.org/10.1186/gm8>
- Collins, P. H. (2015). Intersectionality's definitional dilemmas. *Annual Review of Sociology, 41*, 1–20. <https://doi.org/10.1146/annurev-soc-073014-112142>
- Covington, J. (1995, October). Racial classification in criminology: The reproduction of racialized crime. *Sociological Forum, 10*(4), 547–568. <https://doi.org/10.1007/BF02095768>
- Dache, A., Haywood, J. M., & Mislán, C. (2019). A badge of honor not shame: An AfroLatina theory of Black-imiento for US higher education research. *Journal of Negro Education, 88*(2), 130–145. <https://www.muse.jhu.edu/article/802613>
- Denton, N. A., & Deane, G. D. (2010). Researching race and ethnicity: Methodological issues. In P. H. Collins & J. Solomos (Eds.), *The SAGE handbook of race and ethnic studies*. Sage.
- Ford, K. S. (2019). Observer-identification: A potential threat to the validity of self-identified race and ethnicity. *Educational Researcher, 48*(6), 378–381. <https://doi.org/10.3102/0013189X19860803>
- Ford, K. S., & Holland, M. M. (2020). Identifying “race unknown” postsecondary students in survey data. *Teachers College Record*. <https://www.tcrecord.org/Content.asp?ContentId=23485>
- Ford, K. S., Rosinger, K., & Choi, J. (2022). A product of prestige? “Race unknown” and competitive admissions in the United States. *Policy Futures in Education*. <https://doi.org/10.1177/14782103211043555>
- Fries-Britt, S., George Mwangi, C. A., & Peralta, A. M. (2014). Learning race in a U.S. context: An emergent framework on the perceptions of race among foreign-born students of color. *Journal of Diversity in Higher Education, 7*(1), 1–13. <https://doi.org/10.1037/a0035636>
- Galvez, E., & Muñoz, S. M. (2020). (Re)Imagining anti-colonial notions of ethics in research and practice. *Journal of College Student Development, 61*(6), 781–796. <https://doi.org/10.1353/csd.2020.0075>
- Garcia, J. A. (2017). The Race Project: Researching race in the social sciences researchers, measures, and scope of studies. *Journal of Race, Ethnicity, and Politics, 2*(2), 300–346. <https://doi.org/10.1017/rep.2017.15>
- Garcia, N. M., & Mayorga, O. J. (2018). The threat of unexamined secondary data: A critical race transformative convergent mixed methods. *Race Ethnicity and Education, 21*(2), 231–252. <https://doi.org/10.1080/13613324.2017.1377415>
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Prentice-Hall.
- 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>
- Go, J. (2020). Race, empire, and epistemic exclusion: Or the structures of sociological thought. *Sociological Theory, 38*(2), 79–100. <https://doi.org/10.1177/0735275120926213>
- Gomez, S. L., & Glaser, S. L. (2006). Misclassification of race/ethnicity in a population-based cancer registry (United States). *Cancer Causes Control, 17*(6), 771–781. <https://doi.org/10.1007/s10552-006-0013-y>
- Groman, R., & Ginsburg, J. (2004). Racial and ethnic disparities in health care: A position paper of the American College of Physicians. *Annals of Internal Medicine, 141*(3), 226.
- Hahn, R. A., Truman, B. I., & Barker, N. D. (1996). Identifying ancestry: The reliability of ancestral identification in the United States by self, proxy, interviewer, and funeral director. *Epidemiology, 7*(1), 75–80.
- Haney López, I. (1997). *White by law: The legal construction of race*. New York University Press.
- Harper, S. R. (2012). Race without racism: How higher education researchers minimize racist institutional norms. *Review of Higher Education, 36*(1), 9–29. <https://dx.doi.org/10.1353/rhe.2012.0047>
- 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>
- Hernández, T. K. (2018). *Multiracials and civil rights: Mixed-race stories of discrimination*. New York University Press.
- Hirschman, D., Berrey, E., & Rose-Greenland, F. (2016). Dequantifying diversity: Affirmative action and admissions at the University of Michigan. *Theory and Society, 45*(3), 265–301. <https://doi.org/10.1007/s11186-016-9270-2>
- Hitlin, S., Brown, J. S., & Elder, G. H. (2007). Measuring Latinos: Racial vs. ethnic classification and self-understandings. *Social Forces, 86*(2), 587–611. <https://doi.org/10.1093/sf/86.2.587>
- Howard, J. (2018). That's not something we have to discuss: Interrupting silences about multiracial students in teacher work. *Urban Review, 50*(4), 693–712. <https://doi.org/10.1007/s11256-018-0462-0>
- Howell, J., & Emerson, M. O. (2017). So what “should” we use? Evaluating the impact of five racial measures on markers of social inequality. *Sociology of Race and Ethnicity, 3*(1), 14–30. <https://doi.org/10.1177/2332649216648465>
- Hu-DeHart, E., & López, K. (2008). Asian diasporas in Latin America and the Caribbean: An historical overview. *Afro-Hispanic Review, 27*(1), 9–21. <https://www.jstor.org/stable/23055220>
- Ishizuka, K. L. (2018). *Serve the people: Making Asian America in the long sixties*. Verso Books.
- James, A. (2001). Making sense of race and racial classification. *Race and Society, 4*(2), 235–247. [https://doi.org/10.1016/S1090-9524\(03\)00012-3](https://doi.org/10.1016/S1090-9524(03)00012-3)
- Jang, S. T. (2018). The implications of intersectionality on Southeast Asian female students' educational outcomes in the United States: A critical quantitative intersectionality analysis. *American Educational Research Journal, 55*(6), 1268–1306. <https://doi.org/10.3102/0002831218777225>
- Johnston, M. P., Ozaki, C. C., Pizzolato, J. E., & Chaudhari, P. (2014). Which box(es) do I check? Investigating college students' meanings behind racial identification. *Journal of Student Affairs Research and Practice, 51*(1), 56–68. <https://doi.org/10.1515/jsarp-2014-0005>
- Johnston-Guerrero, M. P. (2017). The (mis)uses of race in research on college students: A systematic review. *JCSCORE, 3*(1), 5–41. <https://doi.org/10.15763/issn.2642-2387.2017.3.1.5-41>

- Kanakamedala, P., & Haga, S. B. (2012). Characterization of clinical study populations by race and ethnicity in the biomedical literature. *Ethnicity & disease*, 22(1), 96.
- Kelly, J. J., Chu, S. Y., Diaz, T., Leary, L. S., Buehler, J. W., The AIDS Mortality Project Group, & The Supplement to HIV/AIDS Surveillance Project Group. (1996). Race/ethnicity misclassification of persons reported with AIDS. *Ethnicity & Health*, 1(1), 87–94. <https://doi.org/10.1080/13557858.1996.9961773>
- Kohli, R., Pizarro, M., & Nevárez, A. (2017). The “new racism” of K–12 schools: Centering critical research on racism. *Review of Research in Education*, 41(1), 182–202. <https://doi.org/10.3102/0091732X16686949>
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (Vol. 111). University of Chicago Press.
- Ladson-Billings, G. (2012). Through a glass darkly: The persistence of race in education research & scholarship. *Educational Researcher*, 41(4), 115–120. <https://doi.org/10.3102/0013189X12440743>
- Ladson-Billings, G. (2007). Pushing past the achievement gap: An essay on the language of deficit. *The Journal of Negro Education*, 76(3), 316–323.
- Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher*, 35(7), 3–12. <https://doi.org/10.3102/0013189X035007003>
- Lee, C. (2009). “Race” and “ethnicity” in biomedical research: How do scientists construct and explain differences in health? *Social Science & Medicine*, 68(6), 1183–1190. <https://doi.org/10.1016/j.socscimed.2008.12.036>
- López, N., Vargas, E., Juarez, M., Cacari-Stone, L., & Bettez, S. (2018). What’s your “street race”? Leveraging multidimensional measures of race and intersectionality for examining physical and mental health status among Latinxs. *Sociology of Race and Ethnicity*, 4(1), 49–66. <https://doi.org/10.1177/2332649217708798>
- Ma, I. W. Y., Khan, N. A., Kang, A., Zalunardo, N., & Palepu, A. (2007). Systematic review identified suboptimal reporting and use of race/ethnicity in general medical journals. *Journal of Clinical Epidemiology*, 60(6), 572–578. <https://doi.org/10.1016/j.jclinepi.2006.11.009>
- Maghbouleh, N., Schachter, A., & Flores, R. D. (2022). Middle Eastern and North African Americans may not be perceived, nor perceive themselves, to be White. *Proceedings of the National Academy of Sciences U S A*, 119(7). <https://doi.org/10.1073/pnas.2117940119>
- Marquardt, K. L., & Herrera, Y. M. (2015). Ethnicity as a variable: An assessment of measures and data sets of ethnicity and related identities. *Social Science Quarterly*, 96(3), 689–716. <https://doi.org/10.1111/ssqu.12187>
- Matias, C. E., & Boucher, C. (2021). From critical whiteness studies to a critical study of whiteness: Restoring criticality in critical whiteness studies. *Whiteness and Education*, 8(1), 64–81. <https://doi.org/10.1080/23793406.2021.1993751>
- Megyesi, M. S., Hunt, L. M., & Brody, H. (2011). A critical review of racial/ethnic variables in osteoporosis and bone density research. *Osteoporosis International*, 22(6), 1669–1679. <https://doi.org/10.1007/s00198-010-1503-z>
- Merton, R. K. (1972). Insiders and outsiders: A chapter in the sociology of knowledge. *American Journal of Sociology*, 78(1), 9–47.
- Mills, C. W. (1997). *The racial contract*. Cornell University Press.
- Morning, A. (2011). *The nature of race: How scientists think and teach about human difference*. University of California Press.
- Morning, A. (2009). Toward a sociology of racial conceptualization for the 21st century. *Social Forces*, 87(3), 1167–1192. <https://doi.org/10.1353/sof.0.0169>
- Mukhopadhyay, C. C. (2018). Getting rid of the word “Caucasian.” In M. S. Kimmel & A. L. Ferber (Eds.) *Privilege: A Reader* (pp. 231–236). Routledge.
- National Institutes of Health. (2005). *NIH grants policy statement*. http://grants2.nih.gov/grants/policy/nihgps_2003/
- Noe-Bustamente, L., Mora, L., & Lopez, M. H. (2020, August 11). *About one-in-four U.S. Hispanics have heard of Latinx, but just 3% use it*. Pew Research Center. <https://www.pewresearch.org/hispanic/2020/08/11/about-one-in-four-u-s-hispanics-have-heard-of-latinx-but-just-3-use-it/>
- O’Connor, C., Lewis, A., & Mueller, J. (2007). Researching “Black” educational experiences and outcomes: Theoretical and methodological considerations. *Educational Researcher*, 36(9), 541–552. <https://doi.org/10.3102/0013189X07312661>
- Omi, M., & Winant, H. (2014). *Racial formation in the United States*. Routledge.
- Quinn, D. M. (2020). Experimental effects of “achievement gap” news reporting on viewers’ racial stereotypes, inequality explanations, and inequality prioritization. *Educational Researcher*, 49(7), 482–492. <https://doi.org/10.3102/0013189X20932469>
- Quinn, D. M., & Desruisseaux, T. M. (2022). Replicating and extending effects of “achievement gap” discourse. *Educational Researcher*, 51(7), 496–499. <https://doi.org/10.3102/0013189X221118054>
- Rachul, C., Ouellette, C., & Caulfield, T. (2011). Tracing the use and source of racial terminology in representations of genetic research. *Genetics in Medicine*, 13(4), 314–319. <https://doi.org/10.1097/GIM.0b013e3181f5cf9a>
- Renn, K. A. (2004). *Mixed race students in college: The ecology of race, identity, and community on campus*. SUNY Press.
- Renn, K. A. (2009). Educational policy, politics, and mixed heritage students in the United States. *Journal of Social Issues*, 65(1), 165–183. <https://doi.org/10.1111/j.1540-4560.2008.01593.x>
- Riley, D., Ahmed, P., & Emigh, R. J. (2021). Getting real: Heuristics in sociological knowledge. *Theory and Society*, 50, 315–356. <https://doi.org/10.1007/s11186-020-09418-w>
- Roberts, D. (2011). *Fatal invention: How science, politics, and big business re-create race in the twenty-first century*. New Press/ORIM.
- Rockquemore, K. A., Brunsma, D. L., & Delgado, D. J. (2009). Racing to theory or retheorizing race? Understanding the struggle to build a multiracial identity theory. *Journal of Social Issues*, 65(1), 13–34. <https://doi.org/10.1111/j.1540-4560.2008.01585.x>
- Roth, W. D. (2018). Unsettled identities amid settled classifications? Toward a sociology of racial appraisals. *Ethnic and Racial Studies*, 41(6), 1093–1112. <https://doi.org/10.1080/01419870.2018.1417616>

- Roth, W. D., van Stee, E. G., & Regla-Vargas, A. (2023). Conceptualizations of race: Essentialism and constructivism. *Annual Review of Sociology*, 49, 39–58. <https://doi.org/10.1146/annurev-soc-031021-034017>
- Saini, A. (2019). *Superior: The return of race science*. Beacon Press.
- Salinas, C., Jr. (2020). The complexity of the “x” in Latinx: How Latinx/a/o students relate to, identify with, and understand the term Latinx. *Journal of Hispanic Higher Education*, 19(2), 149–168. <https://doi.org/10.1177/1538192719900382>
- Saperstein, A., & Penner, A. M. (2012). Racial fluidity and inequality in the United States. *American Journal of Sociology*, 118(3), 676–727. <https://doi.org/10.1086/667722>
- Saucier, P. K., & Woods, T. P. (2018). *Conceptual aphasia in black: Displacing racial formation*. Lexington Books.
- Sen, M., & Wasow, O. (2016). Race as a bundle of sticks: Designs that estimate effects of seemingly immutable characteristics. *Annual Review of Political Science*, 19(1), 499–522. <https://doi.org/10.1146/annurev-polisci-032015-010015>
- Sewell, A. A. (2016). The racism-race reification process: A meso-level political economic framework for understanding racial health disparities. *Sociology of Race and Ethnicity*, 2(4), 402–432. <https://doi.org/10.1177/2332649215626936>
- Shanawani, H., Dame, L., Schwartz, D. A., & Cook-Deegan, R. (2006). Non-reporting and inconsistent reporting of race and ethnicity in articles that claim associations among genotype, outcome, and race or ethnicity. *Journal of Medical Ethics*, 32(12), 724–728. <https://doi.org/10.1136/jme.2005.014456>
- Shotton, H., Lowe, S. C., & Waterman, S. J. (2012). *Beyond the asterisk: Understanding native students in higher education*. Stylus.
- Shrikant, N., & Sambaraju, R. (2021). “A police officer shot a Black man”: Racial categorization, racism, and mundane culpability in news reports of police shootings of black people in the United States of America. *British Journal of Social Psychology*, 60(4), 1196–1217. <https://doi.org/10.1111/bjso.12490>
- Slemp, K. (2020). *Latino, Latina, Latin@, Latine, and Latinx: Gender inclusive oral expression in Spanish*. University of Western Ontario Press.
- Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist*, 60(1), 16–26. <https://doi.org/10.1037/0003-066x.60.1.16>
- Stevens, G., Ishizawa, H., & Grbic, D. (2015). Measuring race and ethnicity in the censuses of Australia, Canada, and the United States: Parallels and paradoxes. *Canadian Studies in Population*, 42(1–2), 13–34. <https://doi.org/10.25336/P6PW39>
- Sue, D. W. (2004). Whiteness and ethnocentric monoculturalism: Making the “invisible” visible. *American Psychologist*, 59(8), 761–769. <https://psycnet.apa.org/doi/10.1037/0003-066X.59.8.761>
- Swidler, A., & Ardit, J. (1994). The new sociology of knowledge. *Annual Review of Sociology*, 20, 305–329. <https://doi.org/10.1146/annurev.so.20.080194.001513>
- Tabron, L. A., & Thomas, A. K. (2023). Deeper than wordplay: A systematic review of critical quantitative approaches in educationresearch(2007–2021). *Review of Educational Research*, 93(5), 756–786. <https://doi.org/10.3102/00346543221130017>
- Telles, E. (2014). *Pigmentocracies: Ethnicity, race, and color in Latin America*. University of North Carolina Press
- Turner, E. O., Baker, D. J., & Jabbar, H. (2024). Improving researchers’ capacity to address injustice: An introduction to the special issue. *Educational Evaluation and Policy Analysis*, 46(2), 175–191.
- Viano, S., & Baker, D. J. (2020). How administrative data collection and analysis can better reflect racial and ethnic identities. *Review of Research in Education*, 44(1), 301–331. <https://doi.org/10.3102/01623737241239982>
- Viano, S., Baker, D. J., Ford, K. S., & Johnston-Guerrero, M. P. (2024). A latent class analysis of racial terminology in education research: Patterns of racial classifications in AERA journals. *Educational Researcher*, 53(4), 213–222. <https://doi.org/10.3102/0013189X241227901>
- Wang, H. L. (2024, Mar 28). *Next U.S. Census will have new boxes for “Middle Eastern or North African,” “Latino.”* NPR. <https://www.npr.org/2024/03/28/1237218459/census-race-categories-ethnicity-middle-east-north-africa>
- Welner, K. G., & Carter, P. L. (2013). Achievement gaps arise from opportunity gaps. In *Closing the opportunity gap: What America must do to give every child an even chance* (pp. 1–10). Oxford University Press.
- Williams, D. R. (1999). The monitoring of racial/ethnic status in the USA: Data quality issues. *Ethnicity & Health*, 4(3), 121–137. <https://doi.org/10.1080/13557859998092>
- Zuberi, T. (2001). *Thicker than blood: How racial statistics lie*. University of Minnesota Press.
- Zuberi, T., & Bonilla-Silva, E. (2008). *White logic, White methods: Racism and methodology*. Rowman & Littlefield.

Authors

DOMINIQUE J. BAKER is an associate professor of education and public policy at the University of Delaware. Her research focuses on the way that education policy affects and shapes the access and success of minoritized students in higher education.

KARLY S. FORD is an associate professor in the Departments of Sociology and Education Policy Studies at Pennsylvania State University. Her research focuses on the sociology of higher education and the politics of collecting data on socially constructed categories such as race, gender, and immigration status.

SAMANTHA VIANO is an associate professor of education at George Mason University. Her research focuses on evaluating K–12 education policies and assessing school contexts that predominantly affect minoritized student populations and their teachers, with particular focus on school safety and security, school improvement, and high school graduation, in addition to research methods for studying racial equity.

MARC P. JOHNSTON-GUERRERO is an associate dean and professor of higher education in the Morgridge College of Education at the University of Denver. His research focuses on racial dynamics and multiraciality in higher education and student affairs.