

A latent class analysis of racial terminology in educational research:  
Patterns of racial classifications in AERA journals

Samantha Viano\*  
George Mason University

Dominique J. Baker  
University of Delaware

Karly S. Ford  
Pennsylvania State University

Marc P. Johnston-Guerrero  
University of Denver

\*Corresponding Author

\*\*\*\*\*

ACCEPTED VERSION

Published in *Educational Researcher*

Available here: <https://journals.sagepub.com/doi/10.3102/0013189X241227901>

\*\*\*\*\*

Acknowledgements: We are grateful for the research assistance of Natalie Truong, Lisa Delacruz Combs, Leandra Cate, and Wan Yu for their participation in our article coding and validation process. We thank the anonymous reviewers as well as participants in the 2022 American Educational Research Association annual meeting for constructive feedback.

Samantha Viano, PhD, is an assistant professor of education at George Mason University, 4400 University Dr, MS 4C2, Fairfax, VA 22030; [sviano@gmu.edu](mailto:sviano@gmu.edu). 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, teacher retention, and high school graduation in addition to research methods for studying racial equity.

Dominique J. Baker, PhD, is an associate professor of education and public policy at the University of Delaware, 106 Alison Hall West, Newark, DE 19716; [djbaker@udel.edu](mailto:djbaker@udel.edu). 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, PhD, is an associate professor in the departments of Sociology and Education Policy Studies at Pennsylvania State University, 400 Rackley Building, University Park, PA; [karly@psu.edu](mailto:karly@psu.edu). 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.

Marc P. Johnston-Guerrero, PhD, is an associate dean and associate professor in the Morgridge College of Education at the University of Denver, 1999 East Evans Avenue, Denver, CO 80210; [marc.guerrero@du.edu](mailto:marc.guerrero@du.edu). His research focuses on racial dynamics and multiraciality in higher education and student affairs.

### **Abstract**

Educational research commonly uses racial terminology but with little understanding of racial classification patterns across the field. In this study, we surface the use of racial terminology using a census of original research published in American Educational Research Association journals between 2009 and 2019. We do so as an ethical quantification exercise, seeking to further social justice goals by encouraging scholarship on racial terminology in educational research. Using latent class analysis, we identify six classes of research ranging from about a third of articles that use almost no racial terminology to an eighth of articles that use terminology extensively. More recently published articles are more likely to be part of classes with extensive or narrow racial terminology usage and less likely to be in classes that are absent racial terminology. Qualitative research is more likely to use extensive racial terminology, and quantitative research is more likely to be absent of or narrowly use racial terminology. We conclude with recommendations for how future research can build off of these findings to address questions on how to authentically and purposefully use racial terminology in ways that reflect the complex ways people identify themselves to better situate educational research to address racial inequality.

Racial classification is commonly included in educational research, given that race, and more importantly racism, is one of the most powerful organizing structures in society (Golash-Boza, 2016). The inclusion of racial classification in educational research has often been criticized as superficial and promoting deficit-oriented perspectives (Kohli et al., 2017; Ladson-Billings, 2012). For instance, federal policy and research on *achievement gaps* often discuss racial identification in a way that implies that these gaps are caused by belonging to certain racial groups instead of being caused by racism (Harper, 2012; Ladson-Billings, 2006). As Garcia (2017) argues in a meta-analysis of racial measurement in 257 social science surveys, researchers' typical ad-hoc approach to race, "has real restrictions on our understanding and explanations of how racial differences are embedded into societal institutions and can serve only to extend the misrepresentation of social reality or lived experiences" (pp. 329). These criticisms share the concern that haphazard use of racial terminology lessens the potential for educational research to reduce racial inequality.

Research in the United States tends to include what the United States Office of Management and Budget (OMB) calls the minimum categories of "White, Black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander" in addition to "Hispanic, Latino, or Spanish". These definitions are indicative of ways in which racial identities are associated with specific terminology. For instance, those who identify as *Black* might report their racial identity<sup>1</sup> as *African American* just as those who identify as *Hispanic* might prefer the term *Latino*. Educational researchers often use broad racial terminology like *minority* as shorthand for those who identify with non-White racial identities.

---

<sup>1</sup> This wording aims to recognize the racial categories that participants select on forms are rough approximations of their racial identities. While social science researchers technically only have access to racial classifications, these are proxies for racial identity which can be fluid and dependent on the context, historical time, whether self-selected, and more (Viano & Baker, 2020).

While prior studies have examined which racial categories should be included in educational research in order to more accurately represent the range of racial identities in the study setting (see Viano & Baker, 2020), we have little understanding of how educational researchers use racial terminology. This gap means that we do not have a clear understanding of the state of the field and makes it less likely studies will be intentional about using and defining racial terminology. For instance, published research increasingly uses the term *Latinx* (Baker et al., 2022), but public polling indicates only 3-4% of Hispanic Americans identify as *Latinx* (Newport, 2022). Acknowledging that racial category definitions and boundaries are constantly in flux due to shifts in political and social identities, there is power in naming and defining which categories researchers are using to describe populations under study, as racial terminology has salience for the applicability of research to policy and practice (Garcia, 2017). However, educational research does not have the expectation that racial terminology is defined and explained with the kinds of intentionality we expect of other measures.

In this study, we aim to encourage the field to have more introspective analyses of racial terminology in education research by examining the use of broad and specific racial categories in published educational research through the lens of ethical quantification (Espeland & Yung, 2019). Through quantifying educational researchers' racial language, we make visible what can be concealed in research articles (Espeland & Lom, 2015). Converting racial terminology into numbers brings attention to racial language in ways that can then shape opportunity to "facilitate changes in power" (Espeland & Yung, 2019, p. 253). For instance, a growing literature on data disaggregation finds significant heterogeneity within racial categories (e.g., Nguyen et al., 2017). Quantifying this heterogeneity has brought attention and resources to marginalized populations (e.g., Sloan, 2022). Similarly, our analysis on the patterns of racial terminology seeks to motivate

future research on racial classification in educational research that can then help authors, reviewers, and editors think more intentionally about racial categorization. In this way, our quantification of racial language can further social justice goals by clarifying opportunities for critical researchers to continue this type of research, to help the field have racial terminology better match the goals of the research. We take an ethical quantification approach to understanding educational researchers' use of racial terminology through the following research questions:

1. Does the use of specific and broad racial terminology in educational research articles allow us to distinguish articles by their patterns of racial terminology use?
2. To what extent does the prevalence of different patterns of racial terminology in published educational research change over time?
3. To what extent does the prevalence of different patterns of racial terminology in published educational research differ depending on the research methodology used?

We address these questions through examining articles from American Educational Research Association (AERA) journals between 2009 and 2019. We explore the use of broad (e.g., *of color*) and specific (e.g., *Asian American*) racial categorizations. Our study analyzes a census of education research articles in order to provide justification for a research agenda on racial terminology in educational research, but this approach is limited in its ability to provide specific recommendations for those in educational research. We propose this analysis can justify future research that then will be able to lay the groundwork for creating a more racially conscious evidence base that is necessary to further social justice goals (Garcia, 2017; Laughter et al., 2023).

### **Racial Categorization in Educational Research**

There is consensus among scientists that racial difference is created by people in social organizations and is not based on biology or genetics. As such, racial groupings are not fixed and vary by context (Omi & Winant, 2015). The boundaries and meanings of racial groups have shifted over time. In the 19<sup>th</sup> century, Jewish and Irish people were formally excluded from whiteness in the United States, and definitions of Black and Indigenous peoples have been governed by blood quantum and rules of hypodescent that mandate racial classifications based on quantifying ancestry. These boundaries can be based on shared social understandings but can also make little sense to the individuals placed in categories. For example, Middle Eastern people in the United States are categorized as White, though this does not reflect their racialized experiences or how their *street race*<sup>2</sup> is read by other people (López et al., 2018). Part of the challenge with how groups are defined is that, while an individual's *racial self-identity* considers their own personal story and nuanced lived experiences, their *racial category* on a survey is a rough approximation shaped by the institutional contexts and/or the state (Roth, 2016). Racial self-identity is often too complex to fit neatly into schemas of racial categorization imposed by survey items. Consequently, it is important for researchers to be aware that these categories are, at best, approximations of individuals' self-identities because racial categorization shapes opportunity structures and focuses attention on boundaries between groups (Irizarry et al., 2023).

The terminology we use in our writing proxies for what we see as important information to convey about our research. It is common in other fields like epidemiology and political science to examine racial category usage patterns in published research (e.g., Garcia, 2017; Gomez & Glaser, 2006; Megyesi et al., 2011). Recent introspective reviews of educational

---

<sup>2</sup> Street race refers to the racial identity that individuals believe others perceive them as in public. This can often be different than individuals' self-perceived racial identity (López et al., 2018).

research have examined more specific questions on researchers' examinations of racism and use of race in framing research (Garcia, 2017; Harper, 2012; Johnston-Guerrero, 2017; Kohli et al., 2017). These explorations surface what is often overlooked – exposing what might be communicated subconsciously in our field to encourage reflection so that future writing will be more purposeful and precise. In this critical study, we make visible one facet of how educational researchers use language around racial categorization. Bringing attention to racial terminology has the potential to increase the intentionality of educational researchers in what language they use and definitions they provide for racial categorization, better aligning language with researchers' social justice goals (Espeland & Yung, 2019).

## Methods

### Data

We examine publications from journals focused on original research published by AERA including *AERA Open*, *American Educational Research Journal*, *Educational Evaluation and Policy Analysis*, *Educational Researcher*, and *Journal of Educational and Behavioral Statistics* between 2009 and 2019. We limit our search to these journals, as they are highly regarded outlets for disseminating educational research from a variety of disciplines (*AERA Journals Online*, n.d.). Our goal is to focus on the writers' language use in describing their research, so we only review original, empirical articles. Out of 1,623 articles published by these journals in these years, 1,427 included original, empirical research. To learn more detail about how we created this sample, see the online appendix.

The research team<sup>3</sup> coded each article to identify the use of broad racial categorizations and specific racial terminology. We created this coding framework based on prior reviews of

---

<sup>3</sup> We include a positionality statement describing how the authors' identities and experiences impacted data collection, analysis, and interpretation in the online appendix.



racial terminology in peer-reviewed articles in other fields (Ma et al., 2007; Stevens et al., 2015). We augmented this framework as we coded. The framework asked the coder to list the terms associated with the categories White, Black, Asian, American Indian, Hispanic, and Two or more races as well as the broad terminology of *minority*, *underrepresented minority*, and *of color* (in reference to race). For instance, the framework asked the coder to input “Race/ethnicity category(ies) for ‘Hispanic’ used anywhere in the paper” with the options of “Hispanic,” “Latino,” “Latinx,” “N/A,” or “Other...” (which allowed inputting). This exercise of translating racial terminology into binary indicator variables allowed us to quantify complex terminology in order to focus attention on patterns of racial categorization. For more information on the validity of our coding framework and reliability of our coding process, see the online appendix.

### **Analysis**

We use latent class analysis (LCA)<sup>4</sup> to categorize our sample into classes based on racial terminology. LCA is a popular tool for classifying multivariate data. LCA is similar to other data mining tools like cluster analysis, but LCA has the advantage of allowing for testing of models with various numbers of typologies to assess the one that best fits the data (Jung & Wickrama, 2008; Nylund et al., 2007). This approach is well suited for this inquiry because LCA is a multivariate approach to identifying groups given a set of characteristics, helping us to recognize patterns that were previously unclear (Espeland & Yung, 2019; Nylund-Gibson et al., 2023).

The variables that define the latent classes are the racial terminology included in at least 5% of articles.<sup>5</sup> This meant including the top two most frequent options for White (*White*,

---

<sup>4</sup> The use of the term “class” throughout the remainder of this study refers to the classes produced by the LCA procedure. To avoid confusion, we include no discussion of the use of the term “class” in reference to socioeconomic status.

<sup>5</sup> While we initially began testing models with our full list of racial terminology from our coding framework, the LCA models would only converge when all variables had means of at least 0.05 (i.e., they were included in at least 5% of articles). We limit our variables throughout the study to those terms.

*Caucasian*), Black (*Black, African American*), and American Indian or Alaska Native (AIAN; *Native American, American Indian*) as well as the top three options for Asian American Pacific Islander (AAPI; *Asian, Asian American, Pacific Islander*) and Hispanic (*Hispanic, Latino, Gender engaged*<sup>6</sup>) in addition to *multiracial*.<sup>7</sup> We include the broad terminology of *underrepresented minority, of color, and minority*. Each individual racial term is represented by a binary indicator equal to one if the article included that term and zero otherwise. For more information on the selection of our LCA model, see the online appendix.

## Results

We list the prevalence of racial terminology across our corpus in Table 1. The most common terms include *White* (56%), *Black* (47%), *Hispanic* (45%), and *Asian* (39%). The most common broad terminology is *minority* (35%) followed by *of color* (16%) and *underrepresented minority* (11%). Almost a third (31%) of our census included none of these terms. This 31% of studies often included human research participants (but did not report on racial demographics) and were published across all five journals, although half of these articles were published in the *Journal of Educational and Behavioral Statistics*.

### Categorization of Articles Based on Racial Terminology

The first step to establishing the typologies of articles by racial terminology use is to iteratively test LCA models starting with the one class model. As shown in appendix Table A1, the six-class model had the most indicators that it would best fit the data. Articles were relatively well distributed across the classes. The values of each of the variables across classes were logical

---

<sup>6</sup> Gender engaged includes following terms: Latino/a, Latino(a), Latinas(os), Latino/a/x, Latinx, Latina/o, Latinas/Latinos, Latinos/as, and Latinos/Latinas.

<sup>7</sup> The names of these categories of racial terminology come from the OMB Directive 15 which standardized how federal programs in the United States collect and report data on racial classification (Race and Ethnic Standards for Federal Statistics and Administrative Reporting, 1977).

given our understanding of racial terminology from previous literature. As is shown in appendix Table A2, the average latent class probability for each class is above 0.80. Using these fit statistics and theory, we concluded the six-class model best fit patterns in the data.

The marginal predicted probabilities on all racial terminology are shown in Table 2. We display these probabilities graphically in online appendix Figure A1 as is traditional in LCA. The column headings in Table 2 reflect our interpretation of the six classes based on the mean values on the racial terminology. Each of these class names could serve as an adjective before “racial terminology.” For instance, we named the first class *Absent* because these articles, representing 36% of the census, used almost no racial terminology (i.e., *Absent* racial terminology). Apart from the six percent of articles using the term *minority*, 0-3% of articles used any other term.

We named the second class *Sporadic* because these articles had relatively low rates of using racial terminology. Articles had the lowest percentages of several common racial classifications compared to the latter four classes including *White* (60%), *Asian* (10%), *American Indian* (2%), *Hispanic* (33%), and *multiracial* (1%). The proportions of articles using the remainder of the terms tended to be low (even if they were not the lowest) compared to the next four classes.

For the third class, *Narrow*, these articles almost always used the terms *White* (91%), *Black* (99%), and *Hispanic* (84%). They often included *Asian* (51%), but rarely included AIAN terms (3% each) or *multiracial* (3%). *Narrow* racial terminology articles had low rates of using *underrepresented minority* (11%) and *of color* (12%). This pattern led us to conceive of racial terminology use as relatively *Narrow*.

The fourth class, *Widening*, used a wider variety of terms across all categories than the *Narrow* class. All *Widening* articles included the terms *White* and *Black* and almost all included

*Asian* (99%) and *Hispanic* (98%). A sizable percentage of *Widening* articles included AIAN terms (28% *Native American*, 47% *American Indian*) as well as *multiracial* (30%). However, *Widening* articles did not include gender-engaged language for Hispanic categories (0%) and preferred *minority* (56%) over other broader terminology. In these ways, *Widening* articles included a wider set of racial terminology but not necessarily extensive terminology use.

The fifth class is termed *Traditional* because these articles also tended to use the breadth of categories, similar to *Widening*, but often used terminology we might term as being less modern. This is most notably the case since *Traditional* articles have the highest percentage of articles with *Caucasian* (24%) as well as a low percentage of gender-engaged terms for Hispanic (7%). *Traditional* articles also rarely used terms like *underrepresented minority* (6%) and *of color* (10%). *Traditional* articles used expansive terms, but tended to rely more on dated terminology.

The final class is termed *Extensive* because these articles included the widest range of racial terminology. *Extensive* articles often used multiple terms for one category. These articles were the most likely to use broad terms like *of color* (70%). Articles in the *Extensive* class had the highest proportion, compared to the other five classes, in using the terms *Asian American* (42%), *Pacific Islander* (37%), *Native American* (38%), *Latino* (75%), gender-engaged language for Hispanic (31%), and *multiracial* (32%).

### **Trends in Racial Terminology Categorization**

We now address research questions 2 and 3 by fitting a multinomial logistic regression model predicting class membership based on year and methodological paradigm (qualitative, quantitative, or mixed methods), results in online appendix Table A3. For ease of interpretation, we show the predicted probabilities of class membership by paradigm in Table 3 and by year in

Figure 1. As we might expect due to increasing awareness among researchers of racial inequality in education (Baker et al., 2022), articles have lower probability of being in the *Absent* class over time (.44 to .30) and higher odds of being in *Narrow* (.20 to .28), *Widening* (.07 to .09), and *Extensive* (.08 to .16) classes.

We might also hypothesize quantitative research would be less likely to use racial terminology or to do so narrowly based on critiques of this paradigm (Gillborn et al., 2018). We confirm that qualitative articles have lower probability of being in the *Absent* (.26) and *Widening* (.02) classes than quantitative articles (.38 and .10, respectively). Qualitative articles have higher probability of being in the *Extensive* class at .27 compared to .09 for quantitative articles. However, qualitative and mixed methods articles have higher probability of being in the *Sporadic* class (.24 for both) than quantitative articles (.10).

### **Sensitivity Analysis on the Six Classes of Racial Terminology Usage**

We realized one weakness of our coding framework was that it did not necessarily indicate the intensity with which these articles used racial terminology. To better understand the use of racial terminology in each class, we randomly selected a 10% sample of articles from each of the five classes (excluding *Absent*). For this sample, the lead author operationalized *intensity* as the number of times the articles mention each racial categorization term. The lead author also noted when the article only mentioned racial terms when describing the sample or in tables; the logic being that articles that only include racial terms in these kinds of ways indicate superficial descriptions of racial categories.

The results of this sensitivity check are in Table 4. We found our class labels matched well with the indicators of intensity of usage of racial terms. Specifically, Table 4 shows the randomly selected *Sporadic* articles had the lowest intensity of usage of these categories

compared to the remaining classes. For instance, *Sporadic* articles tended to mention racial terms just once or twice and 58% of the sample of these articles only mentioned racial terminology in tables or sample descriptions. Of the randomly selected *Narrow* articles, 46% only mentioned racial categorization in tables/sample description. Nevertheless, the *Narrow* sample tended to have multiple mentions of categories related to *White* (average 13), *Black* (average 13), and *Hispanic* (average 8). Over half of the *Widening* sample only used racial terminology for tables/sample description. While the *Widening* sample used a wide-variety of terms, this was often only in tables where the article would include results for the covariates in multiple tables. We noted that 80% of the *Traditional* sample only used racial terminology in the tables/sample description. Finally, the *Extensive* sample showed the highest intensity of usage across categories with an average of 27 mentions of *Black* and 19 mentions of *White* along with the lowest percentage only mentioning racial categorization in the tables/sample description (24%).

### **Discussion**

In this study, we sought to understand whether we can categorize original research articles in AERA journals based on their use of racial terminology. We found AERA articles grouped into six classes based on racial terminology that ranged from just over a third of articles with no racial terminology (i.e., *Absent*) to roughly one-eighth of articles that extensively used a wide variety of racial terms (i.e., *Extensive*). The second most common class with one-quarter of articles, *Narrow*, included a limited set of racial terms, primarily *White*, *Black*, and *Hispanic*. The other classes in between *Narrow* and *Extensive* used a wider variety of racial terminology, although *Traditional* articles tended to use more outdated terms compared to more modern terminology in the *Widening* articles (e.g., *Caucasian* versus *American Indian*). Quantitative articles were more likely to be in the *Absent* and *Narrow* classes versus qualitative articles which

had higher concentration in the *Sporadic* and *Extensive* classes. We identified a few trends in the proportion of articles in each of these classes over time that indicate being a member of a class that uses more extensive racial terminology became more common. We also found suggestive evidence that articles in the *Extensive* class were more likely to engage with racial classification outside of the sample description and tables.

Quantifying racial classifications and how articles group together based on their racial categorization terms allows us to bring attention to racial terminology (Espeland & Lom, 2015). Differences in class assignment by research paradigm and over time have occurred organically. This trend is an encouraging sign that educational research is incorporating more (potentially) equity-aligned language in articles as indicated by the lower likelihood of empirical research having no racial terminology as well as the higher likelihood of more extensive racial terminology. In defining classes of racial terminology use, this study endeavors to translate this attention into opportunity for future research that can provide more specific recommendations for those writing educational research manuscripts, reviewers, and editors to intentionally align the goals of their research with their use of racial terminology (Espeland & Yung, 2019).

## **Implications**

### ***Further Research on Racial Terminology in Educational Research***

For educational research, this analysis can help provide the framework for a future research agenda that clarifies equity-aligned racial categorizations definitions as well as how racial terminology should be incorporated into analyses for social-justice-oriented research. In particular, our sensitivity analysis illuminated how often racial terminology is used to only describe a study's sample or in tables. Future research should engage with questions to help us understand whether it is appropriate to imply racialization is taking place, by including race as a

covariate, without engaging in why “controlling for race” is necessary/appropriate. We did not analyze our corpus for their authentic engagement with racial identification or racism, but future research that uses different kinds of analytic approaches can seek to understand whether different types of racial-terminology use can proxy for the depth of engagement with racialization in an article. For instance, future research might ask whether articles that use racial terminology that we categorized as *Traditional* also tend to include race superficially. Another similar line of inquiry might address whether what we term as *Extensive* racial terminology is indicative of more racially literate research.

To meet social justice goals, we encourage racial categorization to be an affirmative choice instead of a passive act, and for future research to more clearly identify what this means in practice. Our analysis is limited in assisting in these intentional decisions but future research can build off of our development of a field-level understanding of racial categorization to provide ideas for different models of racial-terminology use that research might further analyze. Specifically, we encourage future research to clarify how racialization should be considered in the analysis (Garcia, 2017; Harper, 2012; Johnston-Guerrero, 2017; Kohli et al., 2017).

While our research design does not allow us to provide specific recommendations for educational researchers, our engagement in this research and related projects (Baker et al., 2022; Ford, 2019; Johnston-Guerrero, 2017; Viano & Baker, 2020) encourages authors to take into consideration the complexity of racial identity and the changing nature of racial terminology. Racial identity is complex, so researchers should consider how their writing addresses this complexity. For instance, authors who only include the term *Hispanic* might do so with the assumption that *Hispanic* and *Latino* are synonymous. However, people who actually identify as *Latino* might not consider themselves *Hispanic* (Martínez & Gonzalez, 2021). To be social-



justice oriented, racial terminology should not reflect assumptions or OMB Directives alone – it should also capture the people behind the data. In other words, authors should reflect on the wording in their data and/or the complexity of how those in the sample/being discussed would identify themselves. Similarly, authors engaged in critical research might also use the term *Latinx* because of its appeal as a gender-engaged term even though it is not grammatically correct in Spanish and is controversial among those who might be classified under that term (Newport, 2022). Our analysis did not determine whether research recognized these kinds of conflicts in their writing (e.g., using the term *Latinx* but discussing its controversial usage), so we encourage future research to examine the extent to which authors recognize when they make decisions determining what racial terminology to use in their work.

Other terminology could be related to how individuals in the data identify but have connotations that are antithetical to social justice. While the *Traditional* group tended to use many terms that continue to have resonance today, these articles had high rates of using the term *Caucasian*, despite this term having racist pseudoscience origins as a marker of skin tone and whiteness-oriented beauty standards (Mukhopadhyay, 2018). Even if some might consider themselves *Caucasian*, we have good reason to discontinue the use of this term. The use of other terms will depend on the context. For instance, in contexts with high populations of both *Asian Americans* and *Pacific Islanders*, *AAPI* would be an accurate term, but this combined terminology might be less appropriate in contexts that are almost exclusively *Asian American*. Similarly, to refer simply to individuals in the data as *Latino* would only be appropriate when the data only describe those who identify as male/as Latino in order to use the term in a gender-engaged manner. In other words, we encourage social justice-oriented researchers to approach racial terminology with the same intentionality of a quasi-experimental study reviewing threats

to internal validity or an ethnography discussing reflexivity. The goal is not to have every article include pre-specified racial terms, as specific discursive practices are not equivalent to more justice-oriented research. Still, we note that, based on past studies focused on the use of race/racism in research (Garcia, 2017; Harper, 2012; Johnston-Guerrero, 2017; Kohli et al., 2017) as well as our findings on the clustering of articles in classes with more sporadic or more extensive racial terminology, intentional and clearly defined racial categorization is more likely to serve education's social justice goals. Understanding the limitations of the current study, we hope this work contributes to a burgeoning body of research on racial categorization in education research.

### ***Educational Research Editors & Editorial Boards***

As part of the social movement spurred by the murder of George Floyd in 2020, AERA released a statement formalizing their commitment to scholarship advancing racial justice along with the announcement of forthcoming special issues on related topics (American Educational Research Association, 2020). Beyond these specific special issues, little has been communicated about specific strategies to support research on race/racism. This contrasts with another prominent social science research organization, the American Psychological Association (APA). After APA released their "Apology to People of Color for APA's Role in Promoting, Perpetuating, and Failing to Challenge Racism, Racial Discrimination, and Human Hierarchy in U.S." in 2021, the APA released official recommendations for how to more equitably include racial identity in research including guidance like "Use precise terminology and describe how and why you are using certain racial and ethnic terms" (Wang & Leath, 2023). We hope that AERA and similar educational research organizations can both reflect on APA's suggestions as

well as support future research that could then be translated into similar guidance for educational researchers.

In making these suggestions, we harken back to our argument that these suggestions should be considered best practice in the same ways we consider high-quality research methods. This is the path taken by APA, and we argue educational research can adopt similar tenets while also expanding on these suggestions based on future research.

### **Limitations and Future Research**

We only observe articles in their published form. Language in peer-reviewed journal articles does not necessarily reflect the preferred terminology of the authors, as articles are revised based on comments from peer reviewers and editors as well as changes made during editing. Our goal is not to call out individual authors but to call attention to trends across articles and patterns in language use that are more or less likely to be indicative of critical research.

Our sensitivity analysis supported that our class descriptions like *Sporadic*, *Extensive*, or *Narrow* matched with how often articles used racial terms and whether racial terminology was used only in tables or to describe the sample. At the same time, we recognize that specific articles in the *Sporadic* or *Narrow* classes might intensively focus on one racial group (e.g., *African American* students). In a scan of all titles in the *Sporadic* class, we identified nine articles (5%) that indicated the article focused on one specific racial group. However, we are not “grading” specific articles (one of the reasons we attempt to describe the classes using nominally scaled descriptors). Our goal is to understand how groups cohere to position educational research to be able to recommend potentially fruitful areas of future research that can ask questions about whether research is authentically and purposefully engaging with racial terminology in ways that reflect the complex ways people identify themselves (Garcia, 2017).

As racial terminology is context-dependent and evolving, future reviews might consider other ways to think about racial terminology. For instance, we did not examine patterns in capitalization of terms like *White* and *Black*. However, recent debates have discussed the political implications of capitalization (see Ewing, 2020). We also note how quickly certain terms have been adopted like *BIPOC* and *Latine*, which did not exist in AERA journals in our sample but have quickly gained popularity over the last few years (Marquez, 2020). Future explorations of racial terminology should be sensitive to these trends and conversations.

By collecting data on a census of original research in AERA journals between 2009 and 2019, we included a wider sample compared to similar studies, which limited our ability to recommend specific terms or patterns of terms as a “gold standard” for critical research. In other words, we can suggest that more extensive use would be preferable to absent racial terminology, but we cannot directly link extensive use with social justice-oriented language. This is in contrast to studies like Harper (2012) which examined studies published between 1999 and 2009 in higher education-focused journals, reviewing 255 articles, and Kohli et al. (2017) which reviewed 186 education articles that analyzed racism. While these previous analyses were, inarguably, more in-depth explorations of language related to racism, our goal was to analyze a broad, comprehensive census. We see our work as building on prior research and hope that it helps guide other scholars to continue expanding our understanding of how education researchers treat race and racism in their peer-reviewed research. Future work could either focus on more specialized journals or a random subset of articles to understand in more detail how race is discussed, potentially implementing methods like critical discourse analysis or a mixed methods design. Since LCA is a data-driven exercise that might not replicate in articles published in educational research journals not published by AERA or AERA journal articles published

outside of this time period, future explorations might find different groupings of articles that could be helpful for further understanding racial terminology use in educational research. These types of explorations would help the field of educational research continue examining how racial terminology has been used and could be potentially changed to further social justice goals and the role of research in addressing racial inequality.

### References

*AERA Journals Online*. (n.d.). American Educational Research Association.

<https://www.aera.net/Publications/Journals>

American Educational Research Association. (2020, September 18). *Research organizations announce joint commitment to advancing scholarly study of racism*.

<https://www.aera.net/Newsroom/Research-Organizations-Announce-Joint-Commitment-to-Advancing-Scholarly-Study-of-Racism>

Baker, D. J., Ford, K. S., Viano, S., & Johnston-Guerrero, M. P. (2022). *Racial category usage in education research: Examining the publications from AERA journals (22–596; EdWorkingPaper)*. Annenberg Institute at Brown University.

<https://doi.org/10.26300/r9dg-kd13>

Espeland, W., & Lom, S. E. (2015). Noticing numbers. In M. Kornberger, L. Justesen, J. Mouritsen, & A. Koed Madsen (Eds.), *Making Things Valuable* (pp. 18–37). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198712282.003.0002>

Espeland, W., & Yung, V. (2019). Ethical dimensions of quantification. *Social Science Information*, 58(2), 238–260. <https://doi.org/10.1177/0539018419851045>

Ewing, E. L. (2020, July 2). I'm a Black scholar who studies race. Here's why I capitalize 'White' [Medium]. *ZORA*. <https://zora.medium.com/im-a-black-scholar-who-studies-race-here-s-why-i-capitalize-white-f94883aa2dd3>

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>

- 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>
- 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>
- Golash-Boza, T. (2016). A critical and comprehensive sociological theory of race and racism. *Sociology of Race and Ethnicity*, 2(2), 129–141. <https://doi.org/10.1177/2332649216632242>
- Gomez, S. L., & Glaser, S. L. (2006). Misclassification of race/ethnicity in a population-based cancer registry (United States). *Cancer Causes & Control*, 17, 771–781. <https://doi.org/10.1007/s10552-006-0013-y>
- Harper, S. R. (2012). Race without racism: How higher education researchers minimize racist institutional norms. *The Review of Higher Education*, 36(1), 9–29. <https://doi.org/10.1353/rhe.2012.0047>
- Irizarry, Y., Monk, E. P., & Cobb, R. J. (2023). Race-shifting in the United States: Latinxs, skin tone, and ethnoracial alignments. *Sociology of Race and Ethnicity*, 9(1), 37–55. <https://doi.org/10.1177/23326492221114813>
- Johnston-Guerrero, M. P. (2017). The (mis) uses of race in research on college students: A systematic review. *JCSCORE*, 3(1), 5–41. <https://www.jstor.org/stable/48644490>
- Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2(1), 302–317. <https://doi.org/10.1111/j.1751-9004.2007.00054.x>

- 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>
- 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>
- 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>
- Laughter, J., Pellegrino, A., Waters, S., & Smith, M. (2023). Toward a framework for critical racial literacy. *Race Ethnicity and Education, 26*(1), 73–93.  
<https://doi.org/10.1080/13613324.2021.1924130>
- 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., 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>
- Marquez, Y. (2020, December 29). Call Me Latine: New online resource helps queer and non-binary Latines define themselves. *Spectrum South - The Voice of the Queer South*.  
<https://www.spectrumsouth.com/call-me-latine/>



- Martínez, D. E., & Gonzalez, K. E. (2021). “Latino” or “Hispanic”? The sociodemographic correlates of panethnic label preferences among U.S. Latinos/Hispanics. *Sociological Perspectives*, 64(3), 365–386. <https://doi.org/10.1177/0731121420950371>
- 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, 1669–1679. <https://doi.org/10.1007/s00198-010-1503-z>
- Mukhopadhyay, C. C. (2018). Getting rid of the word “Caucasian.” In M. S. Kimmel (Ed.), *Privilege: A Reader* (pp. 231–236). Routledge.
- Newport, F. (2022, January 7). *Controversy over the term “Latinx”: Public opinion context*. Gallup: Polling Matters. <https://news.gallup.com/opinion/polling-matters/388532/controversy-term-latinx-public-opinion-context.aspx>
- Nguyen, B. M. D., Alcantar, C. M., Curammeng, E. R., Hernandez, E., Kim, V., Paredes, A. D., Freeman, R., Nguyen, M. H., & Teranishi, R. T. (2017). *The racial heterogeneity project: Implications for educational research, practice, and policy* (p. 46). ACT Center for Equity in Learning. <https://eric.ed.gov/?id=ED583600>
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(4), 535–569. <https://doi.org/10.1080/10705510701575396>
- Nylund-Gibson, K., Garber, A. C., Singh, J., Witkow, M. R., Nishina, A., & Bellmore, A. (2023). The utility of latent class analysis to understand heterogeneity in youth coping strategies: A methodological introduction. *Behavioral Disorders*, 48(2), 106-120. <https://doi.org/10.1177/01987429211067214>

Race and Ethnic Standards for Federal Statistics and Administrative Reporting, Directive No. 15 (1977).

Omi, M., & Winant, H. (2015). *Racial formation in the United States* (Third edition).  
Routledge/Taylor & Francis Group.

Roth, W. D. (2016). The multiple dimensions of race. *Ethnic and Racial Studies*, 39(8), 1310–1338. <https://doi.org/10.1080/01419870.2016.1140793>

Sloan, K. (2022, December 5). More Asian Americans on the federal bench; Progress lacking at Big Law. *Reuters*. <https://www.reuters.com/legal/legalindustry/more-asian-americans-federal-bench-progress-lacking-big-law-2022-12-05/>

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>

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/10.3102/0091732X20903321>

Wang, M., & Leath, S. (2023, November 2). Introducing APA Style Journal Article Reporting Standards for Race, Ethnicity, and Culture. *APA Style Blog*. <https://apastyle.apa.org/blog/race-ethnicity-culture-reporting-standards>

**Table 1**

*Descriptive Statistics*

	Proportion of Articles
<b>White</b>	
White	0.56
Caucasian	0.06
<i>(no terms for White)</i>	<i>0.42</i>
<b>Black</b>	
Black	0.47
African American	0.34
<i>(no terms for Black)</i>	<i>0.41</i>
<b>AAPI</b>	
Asian	0.39
Asian American	0.08
Pacific Islander	0.10
<i>(no terms for AAPI)</i>	<i>0.60</i>
<b>AIAN</b>	
Native American	0.10
American Indian	0.09
<i>(no terms for AIAN)</i>	<i>0.83</i>
<b>Hispanic</b>	
Hispanic	0.45
Latino	0.21
Gender-engaged language	0.06
<i>(no terms for Hispanic)</i>	<i>0.45</i>
<b>Two or More</b>	
Multiracial	0.08
<i>(no terms for Two or More)</i>	<i>0.92</i>
<b>Broad Terminology</b>	
Underrepresented minority	0.11
Of color	0.16
Minority	0.35
<i>(no Broad Terminology)</i>	<i>0.58</i>
Number of Articles	1,427

*Note.* Articles can use more than one term within category such that the combination of all of the proportions will often be greater than one. AAPI refers to the category Asian American Pacific Islander. AIAN refers to American Indian or Alaska Native. The Hispanic, Gender-engaged category includes the following terms: Latino/a, Latino(a), Latinas(os), Latino/a/x, Latinx, Latina/o, Latinas/Latinos, Latinos/as, and Latinos/Latinas.

**Table 2**

*Summary of the Six-Class Latent Class Analysis Solution*

	Absent	Sporadic	Narrow	Widening	Traditional	Extensive	All
<b>White</b>							
White	0.03	0.60	0.91	1.00	0.71	0.98	0.56
Caucasian	0.00	0.13	0.02	0.00	0.24	0.13	0.06
<b>Black</b>							
Black	0.00	0.27	0.99	1.00	0.10	0.90	0.47
African American	0.01	0.54	0.30	0.29	0.97	0.78	0.34
<b>AAPI</b>							
Asian	0.01	0.10	0.51	0.99	0.94	0.92	0.39
Asian American	0.00	0.00	0.00	0.10	0.24	0.42	0.08
Pacific Islander	0.00	0.01	0.00	0.45	0.26	0.37	0.10
<b>AIAN</b>							
Native American	0.00	0.04	0.03	0.28	0.22	0.38	0.10
American Indian	0.00	0.02	0.03	0.47	0.16	0.27	0.09
<b>Hispanic</b>							
Hispanic	0.00	0.33	0.84	0.98	0.69	0.61	0.45
Latino	0.00	0.25	0.18	0.14	0.42	0.75	0.21
Gender-engaged language	0.00	0.07	0.02	0.00	0.07	0.31	0.06
<b>Two or More</b>							
Multiracial	0.00	0.01	0.03	0.30	0.13	0.32	0.08
<b>Broad Terminology</b>							
Underrepresented minority	0.01	0.12	0.11	0.15	0.06	0.35	0.11
Of color	0.00	0.17	0.12	0.10	0.10	0.70	0.16
Minority	0.06	0.40	0.48	0.56	0.31	0.78	0.35
Number of Articles	508	185	350	116	95	173	1,427
Percent of Articles	36%	13%	25%	8%	7%	12%	100%

*Note.* AAPI refers to the category Asian American Pacific Islander. AIAN refers to American Indian or Alaska Native.

**Table 3**

*Predicted Probability of Class Membership by Methodological Paradigm*

	Absent	Sporadic	Narrow	Widening	Traditional	Extensive
Qualitative	0.262*** (0.0297)	0.239*** (0.0289)	0.124*** (0.0223)	0.0183* (0.00909)	0.0874*** (0.0191)	0.269*** (0.0298)
Quantitative	0.377*** (0.0143)	0.102*** (0.00897)	0.269*** (0.0131)	0.0955*** (0.00870)	0.0628*** (0.00716)	0.0932*** (0.00859)
Mixed methods	0.298*** (0.0567)	0.242*** (0.0529)	0.237*** (0.0519)	0.0445 (0.0251)	0.0639* (0.0308)	0.115** (0.0383)

*Notes.* Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Table 4**

*Exploration of Intensity of Language Use in Random 10% Sample of Classes (Other Than Absent)*

	Sporadic	Narrow	Widening	Traditional	Extensive
White	2.79	12.77	6.83	2.60	19.29
(median)	(1.00)	(2.00)	(4.50)	(1.00)	(12.00)
Caucasian	0.05	0.03	0.00	1.20	0.18
(median)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Black	1.42	12.86	5.17	0.00	26.59
(median)	(0.00)	(3.00)	(4.50)	(0.00)	(3.00)
African American	1.47	2.31	0.08	3.90	4.77
(median)	(1.00)	(0.00)	(0.00)	(1.50)	(3.00)
Asian	0.05	2.09	2.92	2.80	3.59
(median)	(0.00)	(1.00)	(2.00)	(2.00)	(2.00)
Asian American	0.00	0.00	0.08	0.00	1.94
(median)	(0.00)	(0.00)	(0.00)	(0.00)	(1.00)
Pacific Islander	0.11	0.00	0.67	1.00	0.24
(median)	(0.00)	(0.00)	(0.00)	(0.50)	(0.00)
Native American	0.05	0.00	0.92	0.30	0.65
(median)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
American Indian	0.16	0.00	0.50	0.10	0.41
(median)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Hispanic	1.16	7.91	4.08	12.30	8.18
(median)	(0.00)	(3.00)	(3.50)	(1.50)	(2.00)
Latino	0.42	0.63	0.08	1.40	3.35
(median)	(0.00)	(0.00)	(0.00)	(1.00)	(2.00)
Gender-engaged language	0.21	0.00	0.00	0.00	5.59
(median)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Multiracial	0.00	0.00	1.00	0.10	1.00
(median)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Underrepresented minority	0.37	0.11	0.50	1.20	1.82
(median)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Of color	1.79	0.17	0.08	0.00	7.24
(median)	(0.00)	(0.00)	(0.00)	(0.00)	(1.00)
Minority	3.84	2.40	7.33	0.90	4.12
(median)	(0.00)	(1.00)	(0.00)	(0.00)	(1.00)
Only mentions in tables or sample description	0.58	0.46	0.58	0.80	0.24
	(1.00)	(0.00)	(1.00)	(1.00)	(0.00)
<i>N</i>	19	35	12	10	17

*Notes.* Shows mean values on the first line with median values in parentheses in the second line.

Intensity is defined as the number of times each term is mentioned within article.

**Figure 1**

*Predicted Probability of Class Membership by Publication Year with 95% Confidence Intervals*

