Data Consultations, Racism, and Critiquing Colonialism in Demographic Datasheets

Nina Exner
Virginia Commonwealth University

Et al.

Let us know how access to this document benefits you.

UMass Chan
MEDICAL SCHOOL

Journal of eScience Librarianship
putting the pieces together: theory and practice

Data Consultations, Racism, and Critiquing Colonialism in Demographic Datasheets

Nina Exner[1], Erin Carrillo[1], and Sam A. Leif[2]

[1]Virginia Commonwealth University, Richmond, VA, USA
[2]University of Nevada, Las Vegas, NV, USA

Abstract

Objective: We consider how data librarians can take antiracist action in education and consultations. We attempt to apply QuantCrit thinking, particularly to demographic datasheets.

Methods: We synthesize historical context with modern critical thinking about race and data to examine the origins of current assumptions about data. We then present examples of how racial categories can hide, rather than reveal, racial disparities. Finally, we apply the Model of Domain Learning to explain why data science and data management experts can and should expose experts in subject research to the idea of critically examining demographic data collection.

Abstract Continued

Results: There are good reasons why patrons who are experts in topics other than racism can find it challenging to change habits from Interoperable approaches to race. Nevertheless, the Census categories explicitly say that they have no basis in research or science. Therefore, social justice requires that data librarians should expose researchers to this fact. If possible, data librarians should also consult on alternatives to habitual use of the Census racial categories.

Conclusions: We suggest that many studies are harmed by including race and should remove it entirely. Those studies that are truly examining race should reflect on their research question and seek more relevant racial questions for data collection.

Introduction

Kivel suggests that to combat racism, one must first assume that it is everywhere, every day (Kivel 2017, 137). Various works on race, gender, colonialism, and other forms of systemic bias suggest that we should interrogate everyday systems around us because the systems built by society often invisibly incorporate the prejudices and power dynamics inherited from history (see: D'Ignazio and Klein 2020; Held 2019; Kendi 2019)

Research is not excluded from the need for this inquiry. There has been a call to enact critical approaches to research via alternative ways to conceptualize science and research (Anderson 2004; Jo & Gebru 2020). The QuantCrit approach, in particular, suggests that critical theory approaches should be applied to quantitative and statistical analyses. (Gillborn et al. 2018; Vincent-Ruz 2020). Furthermore, computational analytics and data science have particularly come under scrutiny recently. Several scholars have identified ways that data in complex algorithms and particularly the machine learning context are increasingly incorporating and invisibilizing inequity, exacerbating biases even when it is intended to address poverty and inequality (Benjamin 2019; Eubanks 2018; Noble 2018; O'Neil 2016; Zuberi 2001).

As information professionals in the data space, there are many decisions made in the process of defining, collecting, transforming, analyzing, presenting, and sharing data. The authors, therefore, suggest that the data lifecycle is produced by and interwoven with systems. From terminology to technology, data choices are influenced by past and present research and the organizations and communities that create research. Data are not immutable; we experience many mutable decisions when we discuss our patrons' options for interoperable data element operationalization, reusable error correction approaches, and findable taxonomic and ontologic structures. If data is produced by and interwoven with systems, then it can reflect the biases of those systems. Thus, data can contain invisible racism, or antiracism, or even both at the same time, inherited by the systems that create the data.

Categorization: Why and how

In 1977, the federal government created standards for collecting racial and ethnic data. The federal government does not define race and ethnicity, and acknowledges the ambiguity of the terms (60 Fed. Reg. 44680). Merriam-Webster defines "race" as being "based on physical traits regarded as common among people of shared ancestry" (n.d.) and "ethnic" as related to groups based on "common racial, national, tribal, religious, linguistic, or cultural origin or background" (n.d.). A related term that will also be discussed in this article is "nativity," which is the place of one's birth. The Office of Management and Budget (OMB) created the following racial and ethnic categories: American Indian or Alaska Native, Asian or Pacific Islander, Black or African American, Hispanic, and White. A subsequent 1997 categorization disaggregated Asian-descended

Americans from Pacific Island-descended Americans (62 Fed. Reg. 58782). The U. S. Census Bureau, as a federal agency, is required to follow the OMB categories (U. S. Census Bureau 2020). Before the 1997 U.S. definitions, the history of racial categorization has embedded a wide range of social and political choices. At first European naturalists, in an attempt to combine their social perceptions of the world with the science of biology, described the people of the world as they perceived them: explicitly excluding non-European people based on race (Kraeger 2004; Zuberi 2001). This approach, along with related Eurocentrism efforts such as tracking any perceptible proportion of "Black blood," became represented and embedded in the U.S. Census and in state and local demographic surveys (Pratt et al. 2015; Zuberi 2001). The resulting heterogeneity of racial data lacked interoperability, which led to challenges in making cross-boundary and longitudinal comparisons. This led to especially noticeable problems once laws came into place requiring states to benchmark inequities and make improvements. As activism expanded for many groups such as Americans of South and Central American origins, it became necessary to build better estimates of the scope of socioeconomic challenges (Humes and Hogan, 2009).

Thus, in the mid-20th century, researchers began to study racial inequities with the aim of correcting them. To facilitate this, the OMB issued a policy called Directive No. 15, which defined the categories of racial and ethnic data to be collected and reported on by federal statistical programs. With this policy came the very explicit statement that science was in no way involved in developing these categories:

> These classifications should not be interpreted as being scientific or anthropological in nature, nor should they be viewed as determinants of eligibility for participation in any Federal program. They have been developed in response to needs expressed by both the executive branch and the Congress to provide for the collection and use of compatible, nonduplicated, exchangeable racial and ethnic data by Federal agencies. (Office of Management and Budget 1977, 37)

We can see that these categories were, and continue to be, driven by issues of politics and policy. Further, they continue to be revised based primarily on political realities. For example, in the original 1977 list, Asian or Pacific Islander was one racial category. However, in response to advocacy by Native Hawaiians, in 1997 this was updated to create two new categories: Asian, and Native Hawaiian or Other Pacific Islander. However, this nuance is sometimes lost. Though the federal government enacted this change, the umbrella Asian American Pacific Islander category persists elsewhere, including in state-level health statistics.

Now, some Pacific Islander community leaders are advocating to disaggregate Pacific Islanders from Asian Americans to reflect the reality of the Pacific Islander experience (Ishisaka 2020). Pacific Islander or Pasifika people are those originating or living in Micronesia, Melanesia, and Polynesia. Pacific Islander

communities face significant health and socio-economic disparities compared to other groups. In most measures, such as poverty rates and health outcomes, Native Hawaiian and Pacific Islander (NHPI) people fall behind white Americans as well as behind Asian Americans. These disparities are especially evident with COVID-19. In Washington, the state with the third-highest number of NHPI in the country, Native Hawaiians and Pacific Islanders have a seven times higher rate of COVID-19 cases than white and Asian American people (Pacific Islander Community Association of Washington 2020). Furthermore, even the Asian-American category is deeply problematic, covering as it does an enormous range of genetic and geographic heritages. The aggregation of Asian-American students, for example, hides substantial disparities in educational attainment among these diverse populations (Allen et al. 2019).

More recently, leading up to the 2000 census, groups advocated reclassifying people of Middle Eastern origin from white to their own primary race category. This was unsuccessful in part due to disagreement on whether the category should be Middle Eastern (a geographic designation) or Arab American (an ethnoracial designation). Other efforts to add to the primary racial classification have similarly failed due to various reasons, including difficulty crafting effective persuasive arguments, lack of data showing past and ongoing discrimination, and political clout (Prewitt 2005). On the other hand, debates at the time also led to changes in data collection on multiracial heritage among Americans, which allowed for more self-identity but also brought many questions to light about the concept of Blackness in the United States (Brundsma and Rockquemore 2002).

Thus we can see that politics and policy drive the OMB/Census categories, but those categories also affect research. Although race is often used to study health disparities, studies have shown that even among supposedly-well-established disparities like the hypertension gap between Black and white people, there are more complex interactions with cultural background and nationality than there are with race alone (Best and Chenault 2014). Bilheimer and Klein (2010) also discuss issues in the analysis of health disparities. One challenge is measurement errors in collecting racial and ethnic data. Though self-report is the gold standard, people from different countries of origin may not identify with any of the OMB categories. For example, the Pew Research Center found that 26.2 million single-race Hispanics said they were "some other race," referring to those who wrote in an answer that did not fit in the race categories listed on the census (Lopez, Krogstad, and Passel 2021). Black identity is not a monolithic or simple matter either, as both global African Diasporic studies and U.S.-based discussions of Black identity have shown (Best and Chenault 2014; Hernandez and Murray-Johnson 2015; Sanders Thompson and Akbar 2003; Zeleza 2005).

Self-report can be improved, and guidance gradually pushes towards new data collection recommendations for large-scale survey data. In 2015, the Census Bureau tested two alternative ways to allow people to self-report their race and ethnicity, and recommended a format that allowed respondents to choose from more granular countries of origin under the broader current racial and ethnic

categories, and to select as many categories as apply (Cohn 2017). The Institute of Medicine also recommended rolling up granular ethnicities into the OMB standard categories (Bilheimer and Klein 2010). These are just a few examples of why racial data needs different nuance in research than it does in U.S. policy-driven federal data collection.

Race categorization: Doing harm

One of the authors works with an undergraduate Health Sciences course on Health Disparities. A semester project requires the students to collect health statistics and compare them between racial and ethnic groups across different geographic levels, using various sources, primarily federal. A recurring topic students tend to present concerns the "Hispanic Health Paradox," which states that Hispanic/Latine[1] people live longer and have lower death rates from heart disease, cancer, and many of the other leading causes of death than non-Hispanic white residents. This is despite having social disadvantages, including lower incomes and worse access to health coverage. There are many theories why, including the more robust social networks, healthier eating habits, and lower smoking rates among some Hispanic/Latine[2] groups, particularly newer arrivals (Hostetter and Klein 2018).

However, the over-aggregation of ethnic data may be the actual origin of this contradiction. Aggregating all Hispanic/Latine people as one ethnic group creates inaccurate homogeneity that can appear to be a "paradox." Studies that apply more nuance have suggested that the "Hispanic" designation masks important sources of health disparities among Hispanic/Latine groups. Some sources of disparities appear to be more strongly linked to socioeconomic topics, such as whether Hispanic/Latine people have health insurance. Other sources of disparities appear to vary according to more culturally-relevant variables such as whether a person speaks primarily Spanish or English or what country a person grew up in. The Hispanic/Latine population comes from more than 20 countries with differing experiences and social circumstances, including immigration status. This inaccurate homogeneity makes it hard to pinpoint health indicators, such as high rates of diabetes, liver disease, certain cancers, and poor birth outcomes among some Hispanic/Latine groups (Hostetter and Klein 2018). For example, data from the U.S. Department of Health and Human Services, Office of Minority Health (2019) shows that Puerto Ricans had a higher infant mortality rate in 2017 than Cubans, Mexicans, Central and South Americans, and non-Hispanic whites. Similarly, data from the Centers for Disease Control and Prevention show that health is generally worse among Hispanic/Latine people born in the U.S. compared to those born in other countries (Hostetter and Klein 2018). Finally, a 2015 study examined the relationship between race and Hispanic ethnicity, maternal and child

---

1   While Hispanic and Latino/a/e are often used interchangeably, they are not synonymous. Hispanic refers to people from Spanish-speaking countries (which would exclude Brazil), while Latino/a/e refers to people of Latin American descent living in the US.

2   Because the term "Latinx" is drawn from the English-speaking "-x" suffix and does not fit well with Latin American languages, the authors use the term "Latine" to indicate anyone who identifies with Latin American ancestry and ethnicities regardless of what gender or genders they identify with.

nativity, country of origin, and asthma among non-Hispanic white and Hispanic children across 65 Los Angeles neighborhoods and found that while lifetime asthma prevalence was reported among 9% of children, with no significant differences between Hispanic and non-Hispanic whites overall, Hispanic children of non-Mexican origin reported higher odds of asthma compared to non-Hispanic white children. Further, a protective nativity effect was also observed among children of foreign-born mothers compared to U.S.-born mothers (Camacho-Rivera et al. 2015). These examples clearly demonstrate the harm of examining data from a narrow, race-based lens.

Reflection and inclusion: Conceptualizing change

This discussion naturally leads to the question of how researchers can improve on past habits in their use of racial demographics. The Institute for Healing and Justice suggests that we should avoid using race throughout clinical education, research, and practice (Chadha et al. 2020). This approach is probably the simplest and most effective at raising awareness and preventing harm, particularly when research teams add race as a just-in-case sort of variable. At the opposite end of the spectrum, a study that is deliberately and primarily focused on interoperable disparity surveillance must include appropriate categories but would do well examine the current best practices in all disparity monitoring in order to structure their entire analytical pipeline towards maximizing interoperability with other disparity surveillance studies (Dorsey et al. 2014). If the study uses race for disparity surveillance, then the research question and design should reflect and emphasize that.

Certainly, though, there is a wide range between surveillance-scale disparity monitoring designs and removing race as a variable. There is a spectrum between studies that do not need race and studies that absolutely must have race and must collect it in a specific way to interoperate with previous designs. Interoperability is important to many studies, but researchers need to learn to consider which concepts of race we want to be interoperable with. The classification system of the Census demographic categories is built by the politics and specific inequities in play before the turn of the century. This classification system is valuable for addressing those known inequities and tracing impacts then and changes from that time onward, longitudinally. Thus, to collect data according to the 1977 or 1997 Census categories is to prioritize interoperability within a system defined by the inequities that were identified and defined then and have been inherited since. However, it can never move beyond the categories that were defined at that time. Since those categories carry considerable history of racism, research studies can be improved if researchers take the time to evaluate whether interoperability with those inherited concepts is more valuable to the study than learning to examine what lenses on race are being privileged in each study.

Rectifying the history of racism requires reflecting seriously on the roles of power and history in the concepts of race and how that affects the classification systems being considered (D'Ignazio and Klein 2020; Garcia et al. 2018). Power struggles

that were taking place when the Census categories were established, and the then-new drive to track initiatives that were attempting to overturn previous power forces, were explicitly stated to be the purpose for defining the demographic groups according to policy priorities of the day rather than science (Office of Management and Budget 1977). Thus, history is embedded in them, and their use or overuse deserves a strong critique. Although "critical school" approaches such as critical race theory and intersectional feminist approaches have generally been applied to qualitative inquiry and giving voice to unhear voices about lived experiences, it is time to look at quantitative data as a product of social systems and thus in need of the same level of antiracist critique as anything else (Garcia et al. 2018).

One approach that has arisen to take up this idea is "QuantCrit" (Garcia et al. 2018; Gillborn et al. 2018; Lopez et al. 2018; Vincent-Ruiz 2020). QuantCrit researchers suggest that one possibility is to examine the literature for meaningful covariates that should be examined alongside race and then prioritize intersectionality as a source of disparities (López et al. 2018). The downstream sharing of this kind of data might need close consideration, though; users of the shared data might choose not to take the intersectional approach. Since sharing of detailed sociodemographics generally leads to privacy risks anyway, this kind of dataset might be sharable only as a processed crosstab or perhaps anonymized to only sociodemographic data if there are other meaningful variables that could be examined without the demographic inquiry. Taking an intersectional approach, as QuantCrit theorists have advocated, should be informed by a close examination of the literature for potential covariates.

In fact, a close examination of the literature to understand known sociodemographic impacts on the question is generally a good idea. Thus, if a case for biological forces is to be made, then genetic literature should have produced it. Likewise, if a case for social forces is to be made, then a multiplicity of social forces can be considered based on what is known. Thus, another QuantCrit approach is to speak to affected communities or seek the voices of affected groups in order to be informed by researchers and community members who are more expert in the disparity research than a (non-disparities-focused) subject researcher might be. Unfortunately, the literature is subject to the same social forces that all of research is, and so, the literature is not always the best source of equity-focused information. It might be a place to suggest that consultees start for a deeper understanding and reflection on historical biases in data collection. Ultimately, though, the voices and experiences of the racial group(s) of focus are essential and should be included in research design (Gillborn et al. 2018).

Being informed and guided by the interests of the data producers should be prioritized; ethical guidelines on human research widely advise that people must be treated as having authority over themselves and their data (e.g., the Belmont Report, Department of Health, Education, and Welfare 1979). That could be taken as a broader principle that a peoples should have authority over their data and determine what benefits to prioritize in the use of that data, rather than serving as

a source of resources to be profited from as if researchers were drilling for oil. Prioritizing the needs of the peoples being studied is a good general principle, and some relatively recent discussions have emphasized this in the context of stateless nations. Readers might be familiar with the CARE principles and the #BeFAIRandCARE movement (see Global Indigenous Data Alliance 2019). Taking these principles of Collective benefit, Authority to control, Responsibility, and Ethics into account in all data collection would undoubtedly improve the social impact of research. The CARE principles would be good guides for social justice in and out of the study of indigenous peoples. Understandably, it might be hard to convince researchers to apply these principles to all research. Nevertheless, it is a good direction to try.

Suppose this kind of community engagement is not possible, and removing race as a variable is not appropriate to the research question. In that case, we suggest going back to the idea of reflexive research. Researchers should reflect on the research question. What is the need for race in this case? Would it be better represented by a different grouping of races, or perhaps by a more nuanced question? If social aspects of race are essential, then clarify to the participants that social effects are the key. If social effects of race are the topic of the research question, consider advising consultees to make the social construct of race very clear by specifying "socially assigned race" instead of simply race (Jones et al. 2008), while still allowing for the option of self-identifying as multiracial (Dorsey et al. 2014).

One first step towards antiracist justice in quantitative research is making it habitual to reflect on the place of racial representation in quantitative research. Some researchers will apply a demographic sheet by habit (as the lead author of this article has done in the past) and structure it by rote application of the Census categories (as the lead author of this article has also done). This is an ingrained habit for many social researchers, and no single preferable improvement to the practice has been identified.

Therefore, the first step towards systemic change is for more researchers to internalize that the Census race categories are not based in research and should not be habitually applied to research. A nonexpert in demographics may reasonably assume that the Census categories are based in meaningful origins, and be simply uninformed about the political origins of common racial and ethnic datasheet categories. If a researcher considers it, they may also be convinced by other researchers' ongoing use of the same categories.

Furthermore, racial data is a multilevel, nuanced issue that needs reflection in each study design. There are good, complex reasons why removing rote datasheets will be a difficult first step for any researcher, much less making subsequent steps towards complex study-by-study critique of the roles of race and ethnicity. Domain experts, such as researchers, need to be exposed to and then acclimate to the idea that they may have been taught misinformation about race in social data.

Leading the change: Teaching and consulting on demographic critique

To change researchers' racial data practices, it is necessary to address the misinformation or mistaken assumption that the common colonial data categories in the U.S. are scientifically correct and canonical approaches to studying race and ethnicity. However, as we have shown above, there is not one single best practice to which researchers can easily switch by rote. Thus, helping researchers to change their thinking will require two stages. The first stage of change is addressing the habit and related misinformation that has caused researchers to misconceptualize the nature of existing racial and ethnic categorization practices. The second is learning to appropriately apply relevant expertise from a research design to the secondary topic of demographics.

Data consultations and open workshops provide an opportunity to enact this epistemic change. Consultations are a gentle way to address misinformation. Misinformation can range from a lack of awareness regarding the scientificness (or lack of scientificness) of racial constructs to a belief that the differences do not impact data findings and therefore do not matter. Misinformation correction requires situating the change consonant with the person's existing worldview (Lewandowsky et al. 2012). Discussing the intersection between racial constructs and data collection can frame the conversation from the familiar perspective of scientific inquiry. Similarly, people can be resistant to authoritarian corrections of misinformation (Lewandowsky et al. 2012). It is important then to ponder a non-authoritarian and non-threatening approach to encouraging sustained change. Data consultations provide a non-threatening space to begin to address erroneous race-based beliefs. Generalized workshops for awareness can also provide a non-threatening and less personal space, helping build awareness as one of a suite of ideas being shared for a range of attendees. They can also help to introduce the idea that a process of developing a little expertise in social demographics might be relevant to researchers in non-demographic subjects.

Once a researcher starts to move away from rote use of census-based racial and ethnic categories, they need to start learning about alternatives that work for their study design. Considering this perspective, an acclimation period should be expected. The Model of Domain Learning (MDL), a theory that examines expertise development as a process versus product, explains that one begins in an acclimation phase with a new domain (Alexander 1997). This framework provides a supportive, developmental lens through which we can address racial knowledge domain familiarity while recognizing the existing domain expertise which researchers possess. A new domain may be very close to existing domains of proficiency but not the same as those proficient areas. Selective deep topic knowledge can develop in just those specific areas of situational interest without building to overall expertise. In other words, a researcher may be very knowledgeable about how to define core variables according to standards used by those experts, yet their expertise may be much lower in secondary topics.

If a researcher knows nothing about a secondary topic, a researcher might consult a data librarian about other taxonomies or ontologies with high Interoperability. However, most people are at least passingly familiar with the Census categories, so it is natural to reach for those categories automatically. MDL situates the researcher's competence with the domain of racial forces and demographics as not high unlike their subject expert topic, but neither is it nonexistent like (perhaps) their acclimation with the domain of digital preservation practices. The researcher naturally would automatically apply the domain knowledge they have—that gained from being a user of the Census. They are not cued to seek an outside expert the way they might with other data issues where they have no knowledge. Data librarians can provide that cuing by asking, without judgment, whether the demographic datasheet was developed based on specific goals or added using the default categories. Experts in a subject domain are not expected to be experts in the data science domain; this is where the opportunity arises for data consultations to begin collaboratively building expertise in data collection. We can then use the consultation to familiarize researchers with the idea that Census categories are not grounded in research and see if they become interested in exploring further. Additionally, collaborative consultation can provide valuable support for this endeavor. Data librarians and subject librarians can work side-by-side to help researchers make the bridge from primary expertise to secondary expertise with data. Extending these conversations to other research support structures on campus, such as institutional review boards or faculty senate, can provide a holistic campus approach to raising the overall education and expectations of researchers.

Acclimation moves toward competence as individuals obtain subject-matter knowledge, use surface and deep-processing strategies to interact with the knowledge, and increase their personal association with the domain (Alexander 1997). Hence, data consultations can support an individual's movement through the acclimation phase via continued conversations surrounding data collection and aggregation designs. Over time, data consultations can begin the conversation on how to enact systematic, sustainable change in data collection, aggregation, and reporting.

Confusing, inconsistent, and contradictory usage of racial terms and categories impedes the ability to use data and results in future research (Yudell et al. 2016). Disparate implementations of changes in racial categorization would only further muddy the waters. By the data collection stage, having a thorough understanding of what the researcher is examining can ensure that the data collected is valuable and applicable to the research goals. Clarifying classificatory terminology when reporting results provides readers with a deep understanding of how and why the classification exists. Finally, ensuring data aggregation retains fidelity to these classifications provides for nuanced and accurate future use of reported data. Further, calls for journals to begin publishing guidelines rationalizing the use of classificatory terminology would support this systematic change (Yudell et al. 2016).

Despite the rhetoric surrounding racial constructs, researchers and scientists may struggle due to misinformation or the belief that research is a systematic process with little room for change. The one-on-one nature of data consultations provides a unique opportunity to address the persistence of erroneous beliefs in a targeted, specific way while encouraging and formulating data collection methods that can shape systematic change.

Acknowledgements

Disclosures

The content of this article is based upon a panel presentation at RDAP Summit 2021 titled "Data Consultations, Racism, and Critiquing Colonialism in Demographic Datasheets" available at https://osf.io/3uyzq.

References

Alexander, Patricia A. 1997. "Mapping the Multidimensional Nature of Domain Learning: The Interplay of Cognitive, Motivational, and Strategic Forces." *Advances in Motivation and Achievement* 10: 213–250.

Allen, Walter, Chantal Jones, and Channel McLewis. 2019. "The Problematic Nature of Racial and Ethnic Categories in Higher Education." American Council on Education: Race and Ethnicity in Higher Education. https://www.equityinhighered.org/resources/ideas-and-insights/the-problematic-nature-of-racial-and-ethnic-categories-in-higher-education

Anderson, Joan M. 2004. "Lessons from a Postcolonial-Feminist Perspective: Suffering and a Path to Healing." *Nursing Inquiry* 11(4): 238–246. https://doi.org/10.1111/j.1440-1800.2004.00231.x

Benjamin, Ruha. 2019. *Race after Technology*. Medford, MA: Polity Press.

Bilheimer, L. T., and R. J. Klein. 2010. "Data and measurement issues in the analysis of health disparities." *Health Services Research* 45(5 Pt 2): 1489–1507.
https://doi.org/10.1111/j.1475-6773.2010.01143.x

Brunsma, David L., and Kerry Ann Rockquemore. 2002. "What Does 'Black' Mean? Exploring the Epistemological Stranglehold of Racial Categorization." *Critical Sociology* 28(1–2): 101–121.
https://doi.org/10.1177/08969205020280010801

Camacho-Rivera, Marlene, Ichiro Kawachi, Gary G. Bennett, and S. V. Subramanian. 2015. "Revisiting the Hispanic Health Paradox: The Relative Contributions of Nativity, Country of Origin, and Race/Ethnicity to Childhood Asthma." *Journal of Immigrant and Minority Health* 17(3): 826–833.
https://doi.org/10.1007/s10903-013-9974-6

Chadha, Noor, Madeleine Kane, Bernadette Lim, and Brenly Rowland. n.d. "Towards the Abolition of Biological Race in Medicine: Transforming Clinical Education, Research and Practice." Institute for Healing & Justice in Medicine. Accessed May 18, 2021.
https://www.instituteforhealingandjustice.org/executivesummary

Cohn, D'vera. 2017. "Seeking better data on Hispanics, Census Bureau may change how it asks about race." *Pew Research Center* April 20, 2017.
https://www.pewresearch.org/fact-tank/2017/04/20/seeking-better-data-on-hispanics-census-bureau-may-change-how-it-asks-about-race

Department of Health, Education, and Welfare. 2018. "The Belmont Report." Office for Human Research Protections. Last modified January 15, 2018. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html

D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.

Dorsey, Rashida, Garth Graham, Sherry Glied, David Meyers, Carolyn Clancy, and Howard Koh. 2014. "Implementing Health Reform: Improved Data Collection and the Monitoring of Health Disparities." *Annual Review of Public Health* 35(1): 123–138.
https://doi.org/10.1146/annurev-publhealth-032013-182423

Eubanks, Virginia. 2018. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press

Garcia, Nichole M., Nancy López, and Verónica N. Vélez. 2018. "QuantCrit: Rectifying Quantitative Methods through Critical Race Theory." *Race Ethnicity and Education* 21(2): 149–157.
https://doi.org/10.1080/13613324.2017.1377675

Gillborn, David, Paul Warmington, and Sean Demack. 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

Global Indigenous Data Alliance. 2019. "CARE Principles of Indigenous Data Governance." Global Indigenous Data Alliance. https://www.gida-global.org/care

Held, Mirjam B. E. 2019. "Decolonizing Research Paradigms in the Context of Settler Colonialism: An Unsettling, Mutual, and Collaborative Effort." *International Journal of Qualitative Methods* 18 (January): 1609406918821574. https://doi.org/10.1177/1609406918821574

Hernandez, Kathy-Ann C., and Kayon K. Murray-Johnson. 2015. "Towards a Different Construction of Blackness: Black Immigrant Scholars on Racial Identity Development in the United States." *International Journal of Multicultural Education* 17(2): 53–72.
https://doi.org/10.18251/ijme.v17i2.1050

Hostetter, Martha, and Sarah Klein. n.d. "In Focus: Identifying and Addressing Health Disparities Among Hispanics." The Commonwealth Fund. Accessed May 18, 2021.
https://www.commonwealthfund.org/publications/2018/dec/focus-identifying-and-addressing-health-disparities-among-hispanics

Humes, Karen, and Howard Hogan. 2009. "Measurement of Race and Ethnicity in a Changing, Multicultural America." *Race and Social Problems* 1(3): 111–131.
https://doi.org/10.1007/s12552-009-9011-5

Ishisaka, N. 2020. "Why It's Time to Retire the Term 'Asian Pacific Islander.'" The Seattle Times November 30, 2020. https://www.seattletimes.com/seattle-news/why-its-time-to-retire-the-term-asian-pacific-islander

Jo, Eun, and Timnit Gebru. 2020. "Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning." Proceedings of the 2020 *Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery. 306–316.
https://doi.org/10.1145/3351095.3372829

Jones, Camara Phyllis, Benedict I. Truman, Laurie D. Elam-Evans, Camille A. Jones, Clara Y. Jones, Ruth Jiles, Susan F. Rumisha, and Geraldine S. Perry. 2008. "Using 'Socially Assigned Race' to Probe White Advantages in Health Status." *Ethnicity & Disease* 18(4): 496–504

Kendi, Ibram X. 2019. *How to Be an Antiracist*. New York: One World Press.

Kivel, Paul. 137AD. *Uprooting Racism*. 4th ed. Gabriola Island, BC: New Society Publishers.

Kreager, Philip. 2004. "Objectifying Demographic Identities." In *Categories and Contexts*. Oxford: Oxford University Press. https://doi.org/10.1093/0199270570.003.0002

Lewandowsky, Stephan, Ullrich KH Ecker, Colleen M. Seifert, Norbert Schwarz, and John Cook. 2012. "Misinformation and Its Correction: Continued Influence and Successful Debiasing." *Psychological Science in the Public Interest* 13(3): 106–131.

Lopez, Mark Hugo, Jens Manuel Krogstad, and Jeffrey S. Passel. 2021. "Who Is Hispanic?" *Pew Research Center* September 23, 2021.
https://www.pewresearch.org/fact-tank/2021/09/23/who-is-hispanic

López, Nancy, Christopher Erwin, Melissa Binder, and Mario Javier Chavez. 2018. "Making the Invisible Visible: Advancing Quantitative Methods in Higher Education Using Critical Race Theory and Intersectionality." *Race Ethnicity and Education* 21(2): 180–207.
https://doi.org/10.1080/13613324.2017.1375185

Merriam-Webster.com Dictionary. n.d. "Ethnic." Accessed October 12, 2021,
https://www.merriam-webster.com/dictionary/ethnic

———. n.d. "Race." Accessed October 12, 2021.
https://www.merriam-webster.com/dictionary/race

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press

Office of Management and Budget. 1977. "DIRECTIVE NO. 15 Race and Ethnic Standards for Federal Statistics and Administrative Reporting."
https://transition.fcc.gov/Bureaus/OSEC/library/legislative_histories/1195.pdf

———. 1995. "Standards for the Classification of Federal Data on Race and Ethnicity." 60 Fed. Reg. 44680.

———. 1997. "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity." 62 Fed. Reg. 58782.
https://www.govinfo.gov/content/pkg/FR-1997-10-30/pdf/97-28653.pdf

O'Neil, Cathy. 2016. *Weapons of Math Destruction*. New York: Crown.

Pacific Islander Community Association of Washington. n.d. "Covid 19 Community Response Report." Accessed May 18, 2021. https://www.picawa.org/covid19-communityresponse

Pew Research Center. 2020. "What Census Calls Us."
https://www.pewresearch.org/interactives/what-census-calls-us

Pratt, Beverly M., Lindsay Hixson, and Nicholas A. Jones. 2015. "Infographic: Measuring Race and Ethnicity Across the Decades: 1790–2010." *Washington, DC: US Census Bureau*.
https://www.census.gov/data-tools/demo/race/MREAD_1790_2010.html

Prewitt, Kenneth. 2005. "Racial Classification in America: Where Do We Go from Here?" Daedalus: *Journal of the American Academy of Arts & Sciences* 134(1): 5–17.

Thompson, Vetta L. Sanders, and Maysa Akbar. 2003. "The Understanding of Race and the Construction of African American Identity." *Western Journal of Black Studies* 27(2): 80–88.

U.S. Census Bureau. 2021. "Research to Improve Data on Race and Ethnicity." About the Bureau. 2021. https://www.census.gov/about/our-research/race-ethnicity.html

U.S. Census Bureau. 2020. "About Race." https://www.census.gov/topics/population/race/about.html

U.S. Department of Health and Human Services, Office of Minority Health. 2019. "Infant Mortality and Hispanic Americans." Policy and Data > Minority Population Profiles. https://www.minorityhealth.hhs.gov/omh/browse.aspx?lvl=4&lvlid=68

Vincent-Ruz, Paulette. 2020. "#QuantCrit Resources." Accessed May 18, 2021. https://sites.lsa.umich.edu/pvincentruz/quantcrit-resources

Yudell, Michael, Dorothy Roberts, Rob DeSalle, and Sarah Tishkoff. 2016. "Taking Race out of Human Genetics." *Science* 351(6273): 564–565.

Zeleza, Paul Tiyambe. 2005. "Rewriting the African Diaspora: Beyond the Black Atlantic." *African Affairs* 104(414): 35–68. https://doi.org/10.1093/afraf/adi001

Zuberi, Tukufu. 2001. *Thicker than Blood: How Racial Statistics Lie*. Minneapolis, MN: University of Minnesota Press.