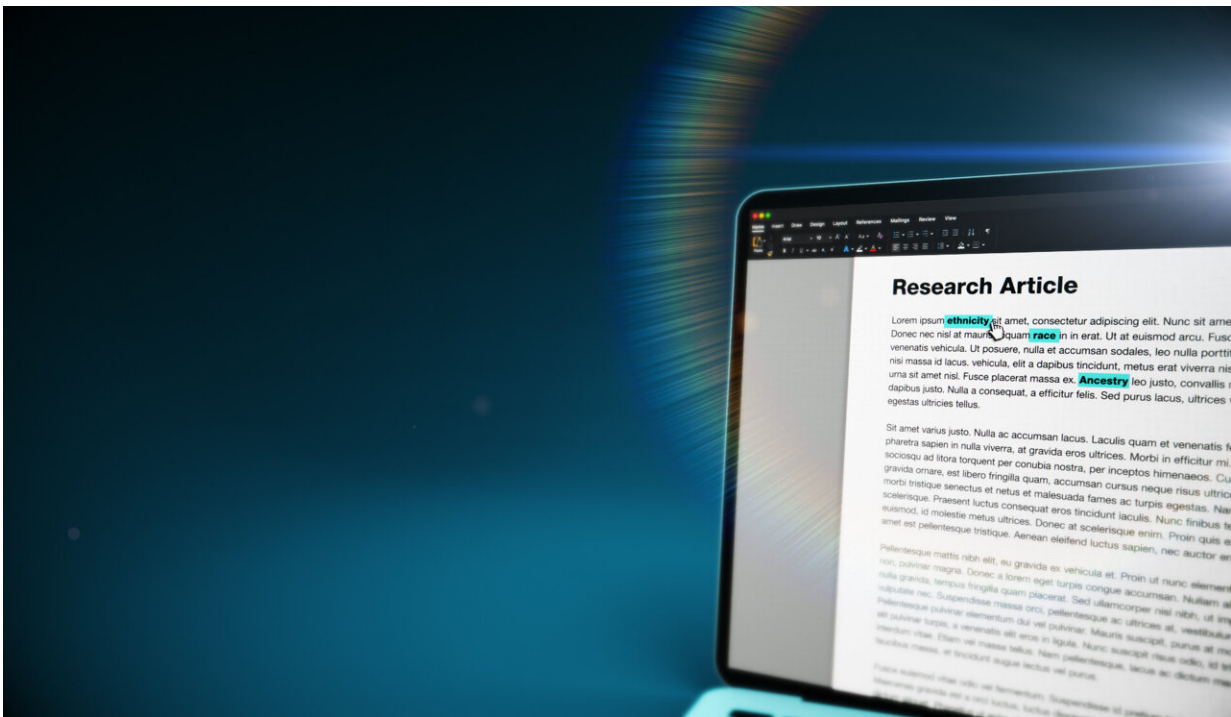


# Language used by researchers to describe human populations has evolved over the last 70 years

December 2 2021



Words that scientists use to describe human populations—such as race, ancestry, and ethnicity—significantly changed from 1949 to 2018. Credit: Ernesto del Aguila III (NHGRI)

National Institutes of Health researchers have found that the words that scientists use to describe human populations—such as race, ancestry, and

ethnicity—significantly changed from 1949 to 2018. Such changes and their timing, along with new descriptors for certain population groups, may be linked to structural racism, social trends, and how people view social constructs such as race.

The study results show that the term "race" is now used less on its own but is used more when paired with "[ethnicity](#)." In addition, the use of "ancestry" and "ethnicity" also has increased. The survey, which was led by researchers at the National Human Genome Research Institute (NHGRI), part of NIH, was published today in the *American Journal of Human Genetics*.

For the field of genetics, the question of what makes a population is a fundamental one. Accurately describing human diversity has a direct impact on our understanding of genomic variation among people, and in turn, how such variation influences our health. Historically, scientists have often incorrectly conceptualized races as distinct biological groups, which has led to health inequities and supported scientific racism.

Many scientists rightly reject the idea of racial and ethnic categories as biological units. Now, in the field of genomics, ancestry, ethnicity and race are used as inexact proxies for genomic ancestry. But scientists still disagree on how these terms should be used and understood.

"Given how ubiquitous these terms are, we wanted to empirically study the historical use of these concepts in the context of genetics and genomics research," said Vence Bonham Jr., J.D., senior author on the study and acting deputy director of NHGRI.

Specifically, the researchers studied the usage of population terms in the 70-year publication history of the *American Journal of Human Genetics*, which is the longest continuously published journal in the field of human genetics. They searched the text of journal articles to identify when

"ancestry," "ethnicity," "race" and other related words began to be used, whether they appeared together, and which terms were slowly used less frequently. Of the 11,635 papers analyzed in the study, 11,360 were [research articles](#), with the rest being award speeches and other communications.

Zhiyong Lu, Ph.D., co-author and a senior investigator in the Intramural Research Program at the National Library of Medicine noted that the use of simple natural language processing programs and robust statistical tests allowed them to analyze tens of thousands of pages easily and find associations between words.

The study's results show that the term "race" appeared in 22% of articles between 1949-58, and declined to 5% between 2009-18; however, in recent years, the term shows up more often when used along with "ethnicity." Conversely, the overall use of "ethnicity" and "ancestry" has increased over time.

Geographic-based terms like "African," "Asian" and "European" are also on the rise. "Hispanic" and "Latina/o/x" were introduced in the journal in 1980 and 1996, respectively.

Of note, descriptors that the authors consider to have negative connotations classically associated with the notion of biological race have declined over the past several decades.

"Some of these shifts could be due to researchers becoming more cognizant of or responsive to the historical and current debates regarding using race in genetics," said Yen Ji Julia Byeon, first author of the study and a doctoral student at Princeton University. "We need ongoing critical reflection as the terminology and concepts used to study human genetic variation continue to shift."

The researchers note that the survey only reveals the quantity of the use of their pre-selected terms rather than the quality. They also acknowledge the limitations of the study given that it was performed on papers in only one journal.

"Future studies can dig deeper into how labels for populations continue to evolve," said Lawrence Brody, Ph.D., co-senior author on the study and director of the NHGRI Division of Genomics and Society. "The goal is to acknowledge our troubled history with [race](#) and build better genomics tools to accurately describe human genomic variation."

**More information:** Yen Ji Julia Byeon et al, Evolving use of ancestry, ethnicity, and race in genetics research—A survey spanning seven decades, *The American Journal of Human Genetics* (2021). [DOI: 10.1016/j.ajhg.2021.10.008](https://doi.org/10.1016/j.ajhg.2021.10.008)

Provided by NIH/National Human Genome Research Institute

Citation: Language used by researchers to describe human populations has evolved over the last 70 years (2021, December 2) retrieved 25 June 2024 from <https://medicalxpress.com/news/2021-12-language-human-populations-evolved-years.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--