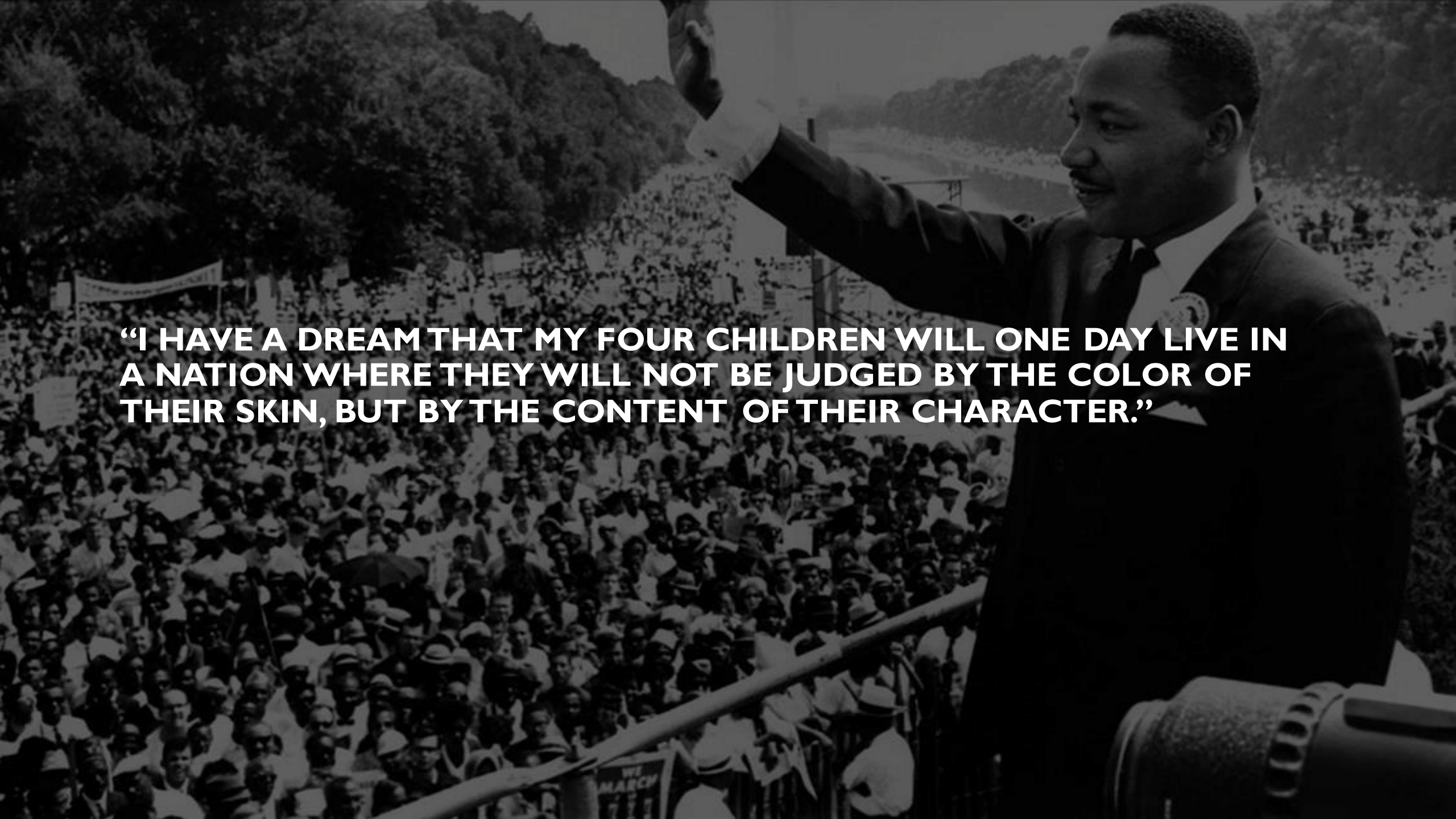
SKIN TONE, COLORISM, AND THE LIMITS OF CENSUS CATEGORIES

Ellis Monk
Professor
Department of Sociology

HARVARD
UNIVERSITY

DISCLOSURES

- **None.**


“I HAVE A DREAM THAT MY FOUR CHILDREN WILL ONE DAY LIVE IN A NATION WHERE THEY WILL NOT BE JUDGED BY THE COLOR OF THEIR SKIN, BUT BY THE CONTENT OF THEIR CHARACTER.”

THE POWER OF METONYMY: RACE AND COLOR

King longed for a day when African Americans, **as a group**, would no longer be held back by systemic and systematic discrimination. **Color**, in this formulation, is a simply a **stand-in** for their membership in a stigmatized ethnoracial **category**.

King's rhetoric relies on a **common linguistic substitution** rooted in the alleged equivalence of belonging to a broad, aggregate ethnoracial category and **a highly conspicuous *marker*** used to ascribe individuals into this broad and heterogeneous ethnoracial category in everyday life.

This slippage, however, was no bar to the efficacy of this famous line – again, the **interchangeability of the terms race and color is so commonplace that the practice continues to this day without drawing much critical attention at all** (e.g. the color line).

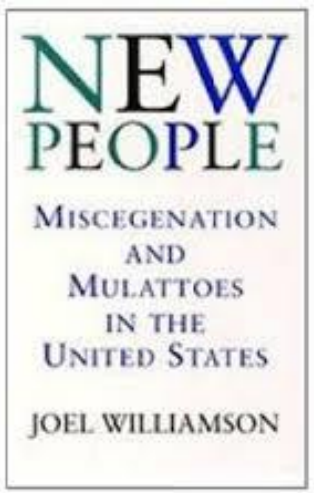
SUBMERGED STRATIFICATION: A BRIEF HISTORY OF COLORISM

- **Color hierarchies originated in slavery**: Lighter-skinned Blacks/mulattos were often favored over darker-skinned/'unmixed' Blacks. After all, these Blacks tended to have direct kinship ties to whites, as it was often sexual violence by whites that initially created this population of lighter-skinned Blacks in the first place.

- **The vast majority of the free Black population was composed of lighter-skinned Blacks and mulattos**. Despite the fact that after Emancipation more opportunities opened up for Blacks of all hues, the substantial social, educational and economic advantages of lighter-skinned Blacks undoubtedly gave these Blacks an undeniably immense head-start in relation to all other Blacks. And evidence suggests lighter-skinned Blacks practiced homogamy and other forms of social closure.

- "**Within the Negro group every possible shade of color between jet Black and creamy white exists**; and variations occur even within the same shade ." 1910 U.S. Census: **21% were 'visibly mulatto.'**

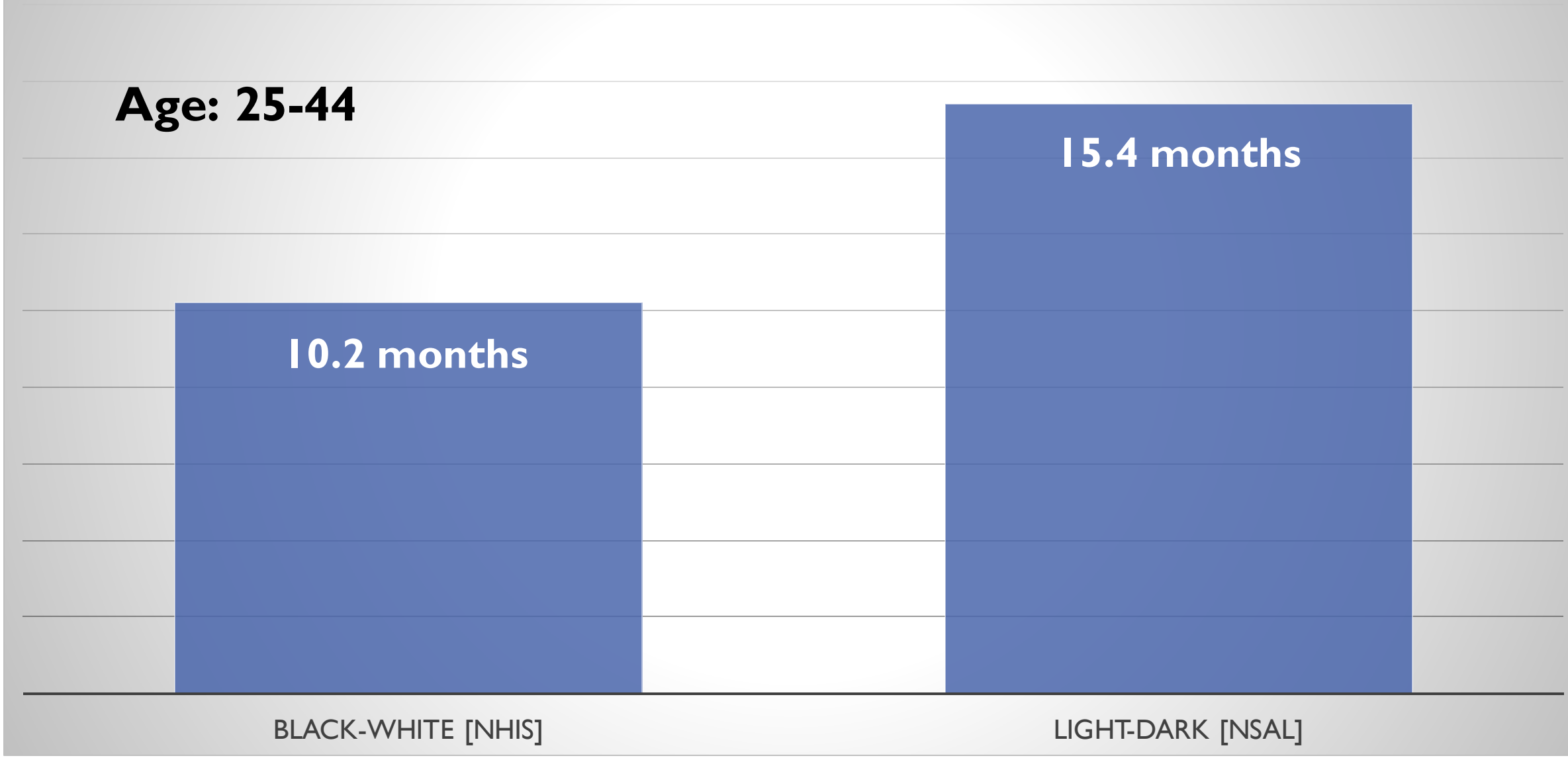
STRATEGIC ESSENTIALISM: "NEW PEOPLE" AND THE ONE DROP RULE



THE PERSISTENT SIGNIFICANCE OF COLOR

- **Census "race/ethnicity" and skin color** are not the same characteristic.
- **There is as much socioeconomic inequality along the color continuum among African Americans as there is between African Americans and whites as a whole** (Keith & Herring 1991; Monk 2014, 2015).
- Evidence shows significant skin tone stratification among **Latinx and Asian Americans** (immigrants and non-immigrants) as well – from earnings to education.
- Colorism is **global**. Research shows colorism has a long history and persists in India (and most of SE Asia), Latin America, throughout Africa, etc.

"Race," Color, and Education in the U.S.



CONCEPTS AND CATEGORIES: THE FOUNDATION OF SOCIAL SCIENCE

Despite ***how fundamental concepts, categories, and categorization are [to research on social inequality]***, most research is conducted ***with little to no reflection on or knowledge of how concepts, categories, and categorization actually work.***

Not integrating contemporary theories of concepts, categories, and categorization makes our research on inequality **poorer and unrealistic**.

The most common measures of difference we use in inequality research, ***State categories***, especially in the way that we tend to use them, stop short of exhausting how social difference relates to inequality.

SOCIAL INEQUALITY:A MATTER OF CATEGORIES?

9. What is Person 1's race? *Mark [X] one or more boxes.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native — *Print name of enrolled or principal tribe.*

- ☐ Asian Indian
- ☐ Chinese
- ☐ Filipino
- ☐ Japanese
- ☐ Korean
- ☐ Vietnamese
- ☐ Other Asian — *Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.*
- ☐ Native Hawaiian
- ☐ Guamanian or Chamorro
- ☐ Samoan
- ☐ Other Pacific Islander — *Print race, for example, Fijian, Tongan, and so on.*

16 What is your ethnic group?

Choose **one** section from A to E, then tick **one** box to best describe your ethnic group or background

A White

- ☐ English/Welsh/Scottish/Northern Irish/British
- ☐ Irish
- ☐ Gypsy or Irish Traveller
- ☐ Any other White background, write in

B Mixed/multiple ethnic groups

- ☐ White and Black Caribbean
- ☐ White and Black African
- ☐ White and Asian
- ☐ Any other Mixed/multiple ethnic background, write in

C Asian/Asian British

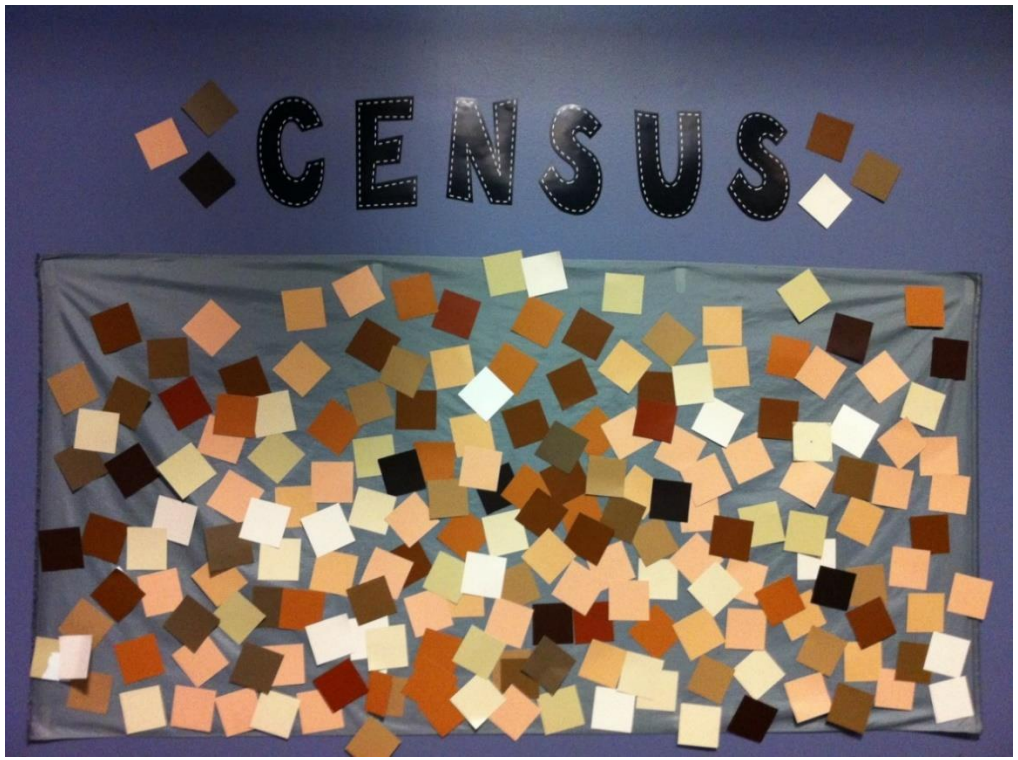
- ☐ Indian
- ☐ Pakistani
- ☐ Bangladeshi
- ☐ Chinese
- ☐ Any other Asian background, write in

D Black/African/Caribbean/Black British

- ☐ African
- ☐ Caribbean
- ☐ Any other Black/African/Caribbean background, write in

E Other ethnic group

- ☐ Arab
- ☐ Any other ethnic group, write in



COMMON FEATURES OF STATE CATEGORIES

Tend to be nominal, [dichotomous], and based on self-identification.

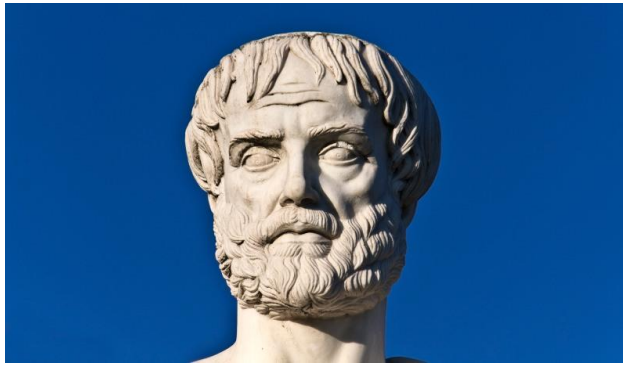
One declares themselves to be a "this" or a "that."

Often taken to be an internally homogeneous "group."

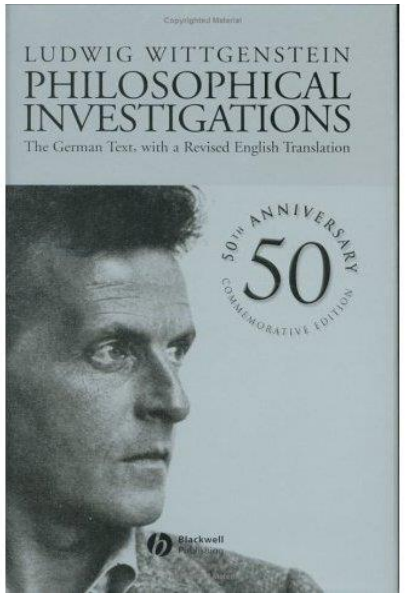
Presumption of **no meaningful differences within a particular category.**

Every member is **equally 'in the category.'**

RETHINKING CONCEPTS AND CATEGORIES: FROM ARISTOTLE TO WITTGENSTEIN.



Aristotelian Theory of Categories (Classical):
Concepts are definitions with necessary and sufficient conditions.
An entity is either 'in' or 'out' of a category.
Every member of a category is an 'equally good' member of that category.



Wittgenstein's Intervention:
Concepts are not definitions with necessary or sufficient conditions.
"Look, don't think…" Concepts are "blurry" and have family resemblances.

THE SHADOW OF THE STATE

"[A] common thread in studies of everyday classification is the recognition that ordinary actors usually have considerable room to maneuver in the ways in which they use even highly institutionalized and powerfully sanctioned categories. They may **adhere nominally to official classificatory schemes while infusing official categories with alternative, unofficial meanings**" (Brubaker et al. 2004: 35).

"[T]he categories used by ordinary people in **everyday interaction often differ substantially from official categories**. The categorized are themselves chronic categorizers; the categories they deploy to make sense of themselves and others **need not match those employed by States, no matter how powerful**..." (Brubaker et al. 2004: 35).

STATE-CENTERED REDUCTION OF SOCIAL REALITY

The State reduces what are **fundamentally continuous social differences** to **crisp, nominal differences** that are more easily handled in their bureaucratic systems (see Bowker and Starr; Bourdieu 1982) and tend to reflect the bases of mass political mobilization that often give rise to said categories.

INFRACATEGORICAL MODEL OF INEQUALITY

(1) Shifting our analytic lens from **mere [nominal] membership in categories to the cues of categories,** subcategories, **and perceived typicality (graded membership).**

This is arguably even more important today than in the past given the relative waning of formal categorical exclusion and increasing contestation over officialized categories (e.g. 'mark one or more' and MENA in the domain of race/ethnicity, categorical flux in the domain of gender and sexuality, etc.)

INFRACATEGORICAL MODEL OF INEQUALITY

(1) **Centering the role of the body** in inequality given its key role in relationally signifying categorical membership (especially ethnoracial and gender categories) and social status (even outside of explicit categorization) via *bodily capital*.

(2) Elevating the analysis of **within-group inequalities** to more equal-footing with between-group inequalities (see Jackson, Chatters, and Taylor 1993; Xie, Killewald, and Near 2016).

(3) Turning our analytic attention to operative, yet **occluded principles of social vision and division** that tend to lack political salience and/or official recognition and don't necessarily refer to "groups" (e.g., colorism, lookism, etc.).

THE COLOR OF HEALTH: EXAMINING HEALTH DISPARITIES BEYOND THE BINARY

TWO PATHWAYS: SOCIAL AND SOCIOTECHNICAL

HEALTH DISPARITIES IN BLACK AND WHITE?

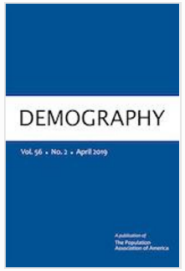
- Health disparities between blacks and whites often remain even after controlling for SES and health behaviors (Das 2013).

- Health disparities *within-race* are often as large or even larger than *between-race* health disparities (Williams & Sternthal 2010).

THE COST OF COLOR: MENTAL AND PHYSICAL HEALTH

- There are significant intraracial gaps in mental and physical health among African Americans (Monk 2015, 2020).

- These intraracial gaps in mental and physical health rival health disparities between Blacks and Whites as a whole (Monk 2015, 2020).



Demography
April 2019, Volume 56, Issue 2, pp 753–762 | Cite as

New Evidence of Skin Color Bias and Health Outcomes Using Sibling Difference Models: A Research Note

Authors | Authors and affiliations

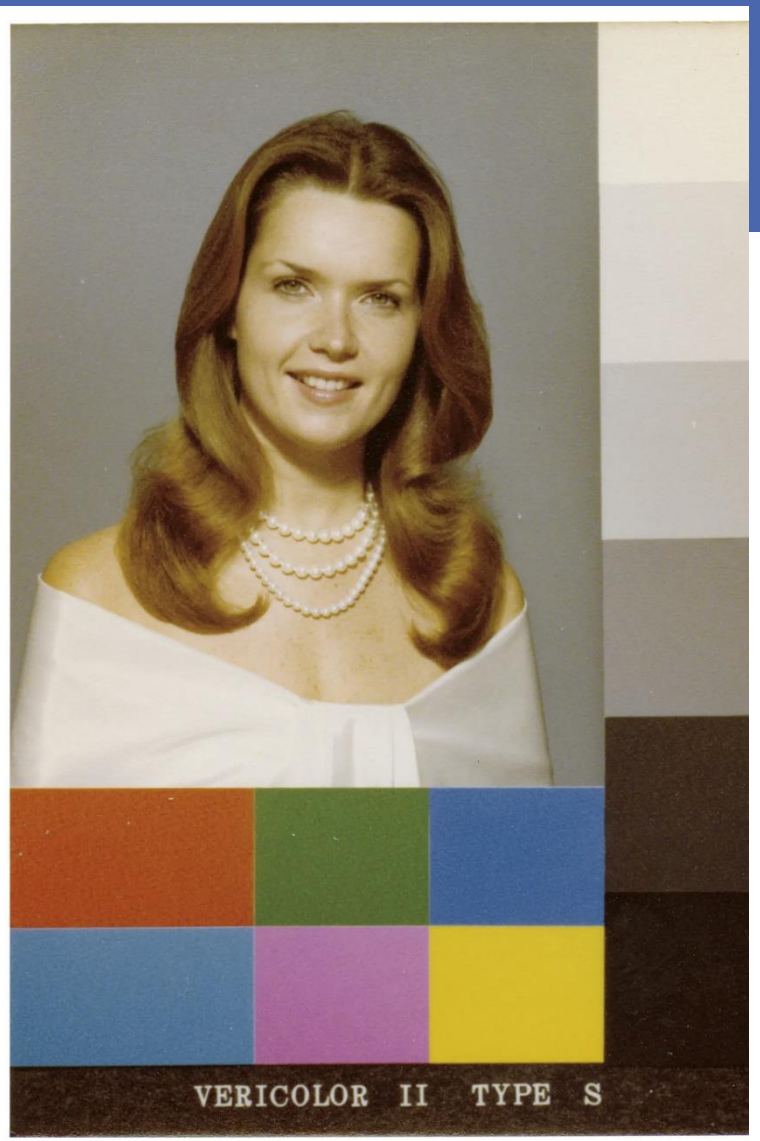
Thomas Laidley, Benjamin Domingue, Piyapat Sinsub, Kathleen Mullan Harris, Dalton Conley

Open Access | Article
First Online: 09 January 2019

Abstract

In this research note, we use data from the National Longitudinal Study of Adolescent to Adult Health (Add Health) to determine whether darker skin tone predicts hypertension among siblings using a family fixed-effects analytic strategy. We find that even after we account for common family background and home environment, body mass index, age, sex, and outdoor activity, darker skin color significantly predicts hypertension incidence among siblings. In a supplementary analysis using newly released genetic data from Add Health, we find no evidence that our results are biased by genetic pleiotropy, whereby differences in alleles among siblings relate to coloration and directly to cardiovascular health simultaneously. These results add to the extant evidence on color biases that are distinct from those based on race alone and that will likely only heighten in importance in an increasingly multiracial environment as categorization becomes more complex.

COLOR-BLIND TECHNOLOGY

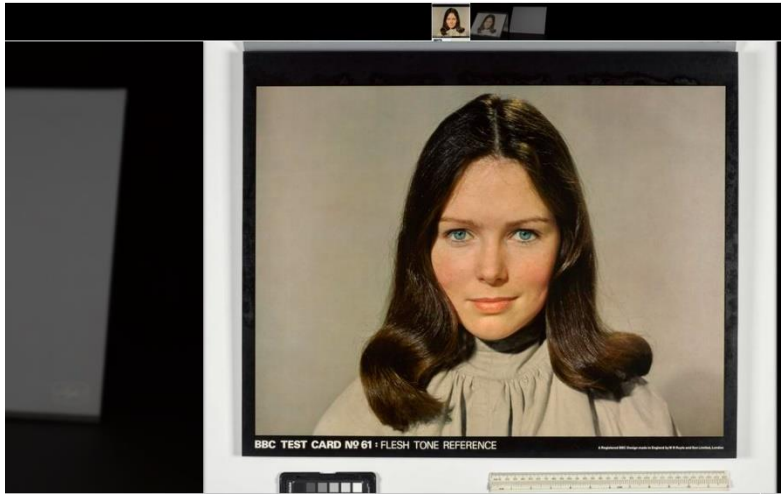


Shirley Card, 1978. Courtesy of Hermann Zschiegner

BBC test card 61, 'Cardboard Kate'

MADE: 1976-1977 in London

MAKER: BBC (Research Department) and W.R. Royle and Son Limited



Autonomous Cars Can't Recognize Pedestrians With Darker Skin Tones

People with darker skin are more at risk of being hit by a self-driving vehicle.

By Jessica Miley
Apr 21, 2021 (Updated: Aug 09, 2021 09:56 EDT)

Social Science & Medicine
Volume 202, April 2018, Pages 38-42

Short communication

Representations of race and skin tone in medical textbook imagery

Patricia Louie [a], Rima Wilkes [b]

Show more

Abstract

Although a large literature has documented racial inequities in health care delivery, there continues to be debate about the potential sources of these inequities. Preliminary research suggests that racial inequities are embedded in the curricular edification of physicians and patients. We investigate this hypothesis by considering whether the race and skin tone depicted in images in textbooks assigned at top medical schools reflects the diversity of the U.S. population. We analyzed 4146 images from *Atlas of Human Anatomy*, *Bates' Guide to Physical Examination & History Taking, Clinically Oriented Anatomy,* and *Gray's Anatomy for Students* by coding race (White, Black, and Person of Color) and skin tone (light, medium, and dark) at the textbook, chapter, and topic level. While the textbooks approximate the racial distribution of the U.S. population - 62.5% White, 20.4% Black, and 17.0% Person of Color - the skin tones represented - 74.5% light, 21% medium, and 4.5% dark - overrepresent light skin tone and underrepresent dark skin tone. There is also an absence of skin tone diversity at the chapter and topic level. Even though medical texts often have overall proportional racial representation this is not the case for skin tone. Furthermore, racial minorities are still often absent at the topic level. These omissions may provide one route through which bias enters medical treatment.

Arthritis Care & Research
American College of Rheumatology

Brief Report

Representation of Skin Colors in Images of Patients With Lupus

Amaad Rana, Abbey Witt, Heather Jones, Muithi Mwanthi, Jacob Murray, Lisa Zickuhr

First published: 31 May 2021 | https://doi.org/10.1002/acr.24712 | Citations: 2

No potential conflicts of interest relevant to this article were reported.

Read the full text > PDF TOOLS SHARE

Abstract

Objective

Lupus presents earlier and more severely among patients with skin of color (SOC), and this population experiences worse outcomes. Providers rely on medical education materials when developing skills to care for patients, yet these resources historically underrepresent patients with SOC and marginalize vulnerable populations. In this study, we investigated if this publication bias extends to images depicting patients with lupus.

AAD JAAD Journal of the American Academy of Dermatology

RESEARCH LETTER | VOLUME 84, ISSUE 1, P194-196, JANUARY 2021

Skin color in dermatology textbooks: An updated evaluation and analysis

Ademide Adelekun, BA • Ginikanwa Onyekaba, BS • Jules B. Lipoff, MD

Published: April 23, 2020 • DOI: https://doi.org/10.1016/j.jaad.2020.04.084 • Check for updates

References

Article info

Linked Article

To the Editor: The US population is becoming more racially and ethnically diverse, yet medical textbooks have not reflected this change. A study of general medicine textbooks showed minimal skin type diversity, with 4.5% of images showing dark skin. [1] In dermatology, visual diagnosis and pattern recognition are affected by the background skin type. Previous work showed limited representation of skin of color in dermatology textbooks. [2] We aimed to update and evaluate any changes in the representation of skin phototypes in core dermatology texts used to educate trainees, dermatologists, and generalists.

PULSE OXIMETRY REGULATIONS

- According to current FDA regulations, **at least 2 volunteers or 15% of the study group**, whichever is higher, should be "darkly-pigmented."

- **Who is "darkly-pigmented"?**

- Fitzpatrick Scale designed in 1975 by Dr. Thomas Fitzpatrick (Harvard Medical School).
- Intended to categorize how skin reacts to UV/tanning during phototherapy. Inconsistent even when used as intended.
- Adopted as gold-standard for skin tone classification, **despite not being designed for this purpose**. Sometimes is used as a **proxy for race**.
- Originally only 4 tones were represented. 10 years later, 2 more were added.

FITZPATRICK SCALE

Studies find, for example, that the Fitzpatrick Scale "**excludes the majority of Blacks and yields data that overestimate Black population prevalence of type IV skin**" (Pichon et al. 2010); and that the "**[Fitzpatrick Scale] provides a restricted range of options for people with darker skin tones that do not capture variations in their skin color**." It is also worth noting that some studies find that the Fitzpatrick Scale performs poorly, **even when used as intended** (see above), especially on ethnoracial minorities (see Eilers et al. 2013; He et al. 2014).

Ironically, given its selection of skin tones, which lives in a very restricted 'intermediate' zone, it may be simultaneously **too dark** for many lighter-skinned people and **not dark enough** for many darker-skinned people.

MONK SKIN TONE SCALE

- Intended to be an **easy to use (e.g., an optimal number of choices), reliable, and cost-effective** means of measuring skin tone.

- The main way the Monk Scale mitigates biases relative to prior visual scales is by including ***a wider range of carefully selected skin tones to better represent the dynamic range of skin tones*** we see in the United States (and beyond).

- Color selection based on extensive fieldwork in the U.S. & Brazil, computer software that creates facial stimuli for social psychological experiments (i.e., skin reflectance spectrum scores), maps of the distribution of UV exposure and human skin tone around the world.

- Validated through cognitive interviewing (NIA/NIH funded research with NSHAP) and nationally-representative surveys. **More inclusive** than the Fitzpatrick Scale and just as easy to use.

Which Skin Tone Measures are the Most Inclusive? An Investigation of Skin Tone Measures for Artificial Intelligence

Courtney M. Heldreth[1]*

Google Research, Seattle, cheldreth@google.com

Ellis P. Monk

Harvard University, emonk@fas.harvard.edu

Alan T. Clark

The Value Engineers, alan.clark@thevalueengineer.com

Susanna Ricco

Google Research, Seattle, ricco@google.com

Xango Eyée

Google Research, Seattle, xango@google.com

Skin tone plays a critical role in artificial intelligence (AI). However, many algorithms have exhibited unfair bias against people with darker skin tones. One reason this occurs is a poor understanding of how well the scales we use to measure and account for skin tone in AI actually represent the variation of skin tones in people affected by these systems. To address this, we conducted a survey with 2,214 people in the United States to compare three skin tone scales: The Fitzpatrick 6-point scale, Rihanna's Fenty™ Beauty 40-point skin tone palette, and a newly developed Monk 10-point scale from the social sciences. We find that the Fitzpatrick scale is perceived to be less inclusive than the Fenty and Monk skin tone scales, and this was especially true for people from historically marginalized communities (i.e., people with darker skin tones, BIPOCs, and women). We also find no statistically meaningful differences in perceived representation across the Monk skin tone scale and the Fenty Beauty palette. We discuss the ways in which our findings can advance the understanding of skin tone in both the social science and machine learning communities.

Consensus and Subjectivity of Skin Tone Annotation for ML Fairness

Candice Schumann
Google
United States
cschumann@google.com

Gbolahan O. Olanubi
Google
United States
femio@google.com

Auriel Wright
Google
United States
aurielwright@google.com

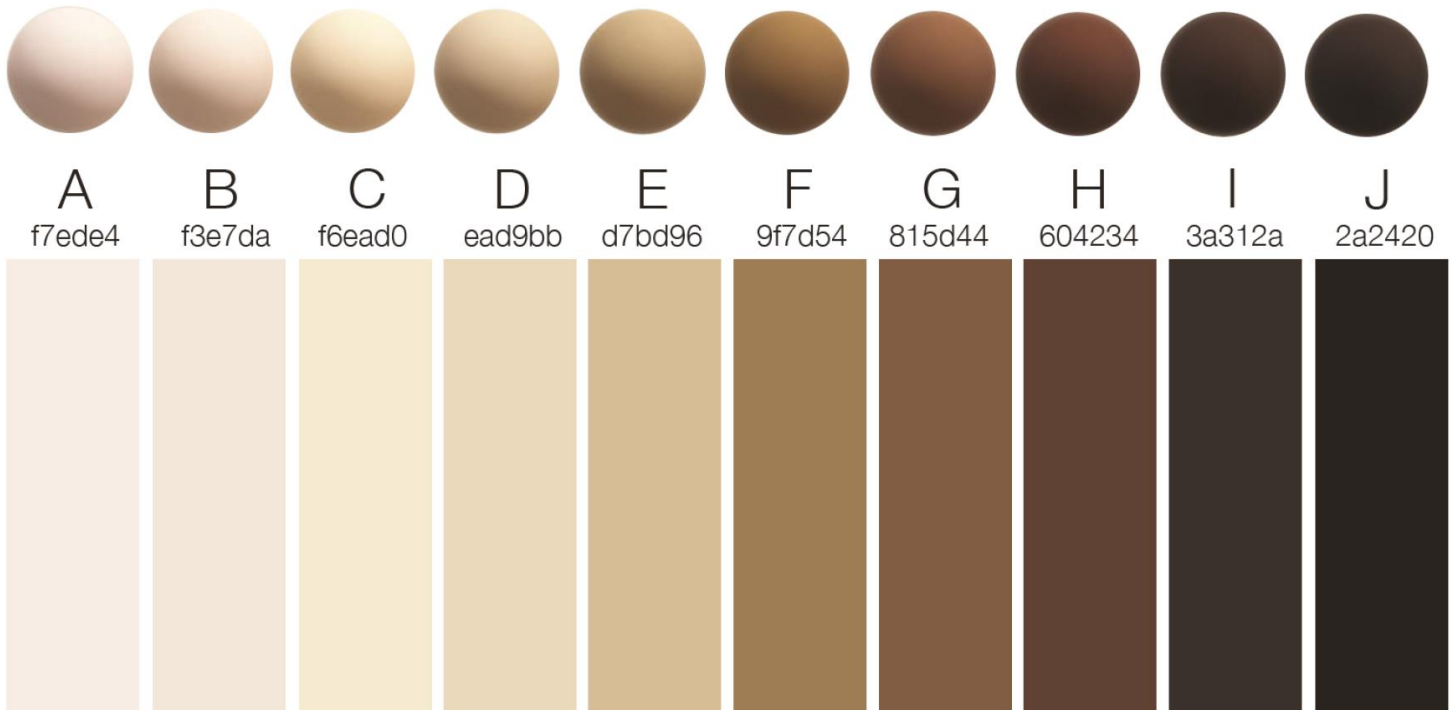
Ellis Monk, Jr.
Harvard University*
United States
emonk@fas.harvard.edu

Courtney Heldreth
Google
United States
cheldreth@google.com

Susanna Ricco
Google
United States
ricco@google.com

Monk Scale

How to use: This scale is intended to be used to classify human skin color. Human skin color is immensely variable and complex, but scales, such as this one, are often used as a pragmatic and cost-effective means of approximately measuring common differences in skin color among human beings. Print this page using color accurate printing. Once printed, consider using a hole punch to punch a hole in the center of each circle and the center of each rectangle. Hold the printed card over the skin at the site you wish to describe. Move the card to find the color that matched most closely and note the category letter (e.g. 'E'). The circles or squares can be used and you should note in your methods which you have chosen. As you may notice the circles have more shades per category, which better resemble human skin and may aid in finding the closest match.



Google @Google

In partnership with Dr. Ellis Monk, we're releasing a new skin tone scale designed to be more inclusive of the full spectrum of skin tones — the next step in our commitment to image equity and improving representation across our products. #GoogleIO

Monk Skin Tone Scale — #GoogleIO

AI

Improving skin tone representation across Google

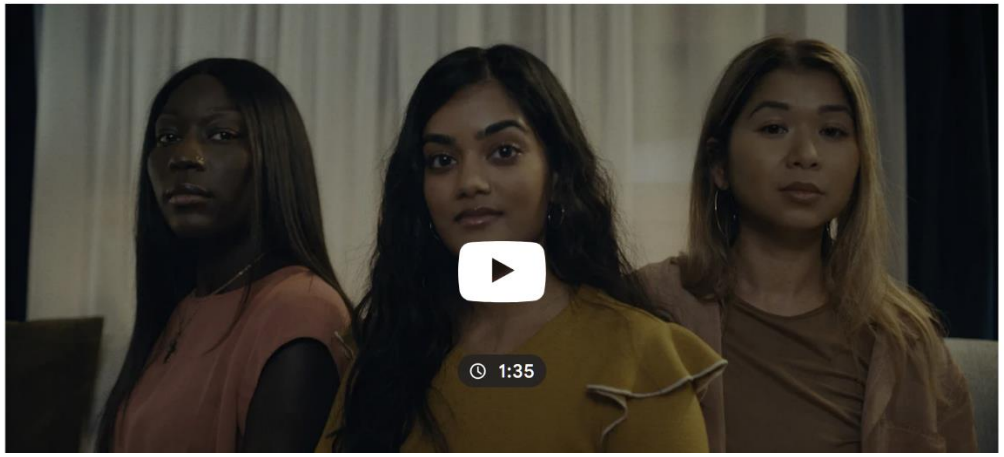
May 11, 2022 · 5 min read

Tulsee Doshi
Head of Product, Responsible AI

Share

This scale was designed to be easy-to-use for development and evaluation of technology while representing a broader range of skin tones. In fact, our research found that amongst participants in the U.S., people found the Monk Skin Tone Scale to be more representative of their skin tones compared to the current tech industry standard. This was especially true for people with darker skin tones.

"In our research, we found that a lot of the time people feel they're lumped into racial categories, but there's all this heterogeneity with ethnic and racial categories," Dr. Monk says. "And many methods of categorization, including past skin tone scales, don't pay attention to this diversity. That's where a lack of representation can happen...we need to fine-tune the way we measure things, so people feel represented."



Improving skin tone representation in Google Photos

We'll also be using the MST Scale to improve Google Photos. Last year, we introduced an improvement to our auto enhance feature in partnership with professional image makers. Now we're launching a new set of Real Tone filters that are designed to work well across skin tones and evaluated using the MST Scale. We worked with a diverse range of renowned image makers, like Kennedi Carter and Joshua Kissi, who are celebrated for beautiful and accurate depictions of their subjects, to evaluate, test and build these filters. These new Real Tone filters allow you to choose from a wider assortment of looks and find one that reflects your style. Real Tone filters will be rolling out on Google Photos across Android, iOS and Web in the coming weeks.

Meta AI Research Publication

RESEARCH

Introducing Casual Conversations v2: A more inclusive dataset to measure fairness

March 9, 2023

Casual Conversations v2: Designing a large consent-driven dataset to measure algorithmic bias and robustness

Abstract

Developing robust and fair AI systems requires datasets with comprehensive set of labels that can help ensure the validity and legitimacy of relevant measurements. Recent efforts, therefore, focus on collecting person-related datasets that have carefully selected labels, including sensitive characteristics, and consent forms in place to use those attributes for model testing and development. Responsible data collection involves several stages, including but not limited to determining use-case scenarios, selecting categories (annotations) such that the data are fit for the purpose of measuring algorithmic bias for subgroups and most importantly ensure that the selected categories/subcategories are robust to regional diversities and inclusive of as many subgroups as possible.

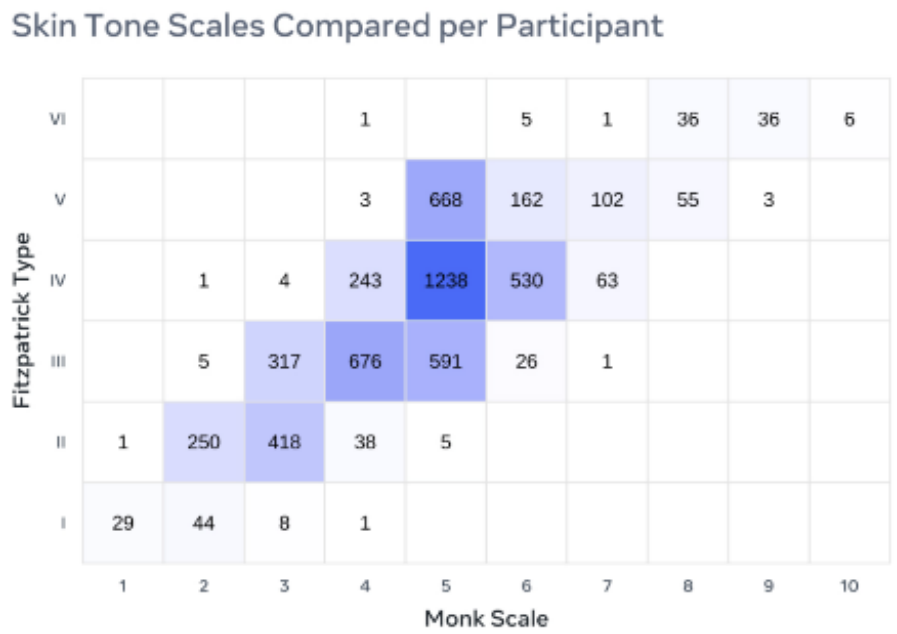
Download Paper

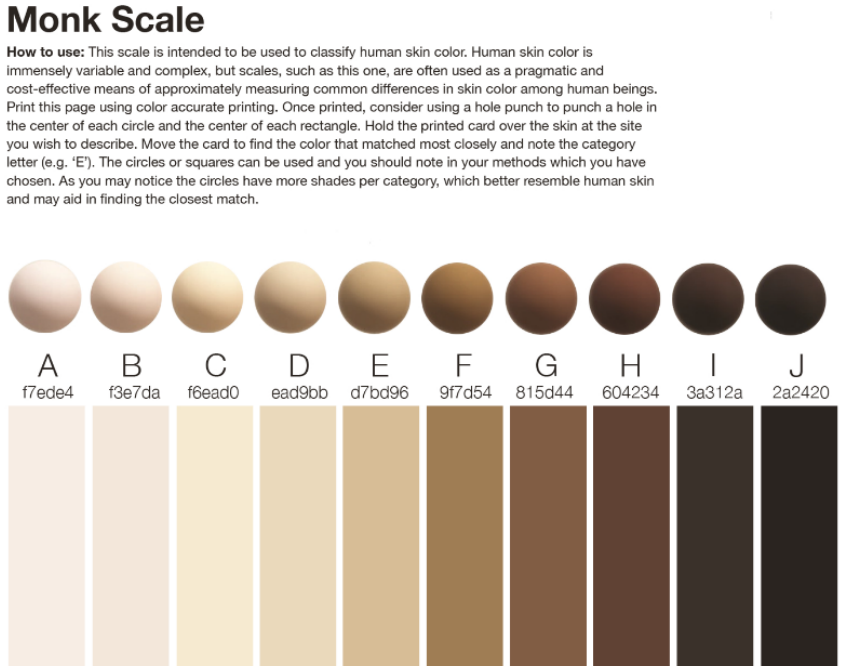
Copy PDF URL

By: Caner Hazirbas, Yejin Bang, Tiezheng Yu, Parisa Assar, Bilal Porgali, Vítor Albiero, Stefan Hermanek, Jacqueline Pan,

While the introduction of recording participants from multiple geographies provided a new set of logistical challenges and opportunities, it also added further complexity in identifying categories relevant to even more diverse communities. With increasing concerns over the performance of AI systems across different skin tone scales, we decided to leverage two different scales for skin tone annotation. The first is the six-tone Fitzpatrick scale, the most commonly used numerical classification scheme for skin tone due to its simplicity and widespread use. The second is the 10-tone Skin Tone scale, which was introduced by Google and is used in its search and photo services. Including both scales in Casual Conversations v2 provides a clearer comparison with previous works that use the Fitzpatrick scale while also enabling measurement based on the more inclusive Monk scale.

Studies find, for example, that the Fitzpatrick Scale **"excludes the majority of Blacks and yields data that overestimate Black population prevalence of type IV skin"** (Pichon et al. 2010);





CONCLUDING THOUGHTS

LOOKING BEYOND THE CENSUS

- Census race/ethnicity and skin tone are NOT the same characteristic.
- There is considerable heterogeneity in skin tone and other phenotypical features associated with race/ethnicity within and across Census race/ethnicity categories.
- Evidence shows, globally, that skin tone is significantly associated with education, earnings, employment, health, and more in many countries around the world.
- While **much attention is often given to race/ethnicity and racism, relatively little attention is paid to color and colorism**, even though it has massive effects on inequality all around the world.

SOCIAL AND SOCIOTECHNICAL COLORISM

- Findings from our research show that the Monk Skin Tone Scale is as easy to use as the Fitzpatrick Scale, while being significantly **more representative and inclusive;** and that there are **high levels of consensus**, globally, using expert and crowdsourced (non-expert) annotators.

- **'Subjective' and 'objective'** measures of skin tone **are both important to collect**.

- Skin tone impacts the probability of diagnosis in dermatology and other ailments (e.g., wounds, AI applications to diagnose various diseases, etc.). **Lack of proper representation of darker skin tones** is at the center of these disparities.

- Different measures may help tap into different mechanisms – **social and/or sociotechnical** - through which skin tone may produce the inaccuracies researchers find with pulse oximeters (e.g., **social determinants of health**). There is still much to learn about the role of skin tone in pulse oximetry.

THANK YOU

Ellis Monk

Professor

Department of Sociology

HARVARD
UNIVERSITY