Why, when, and from whom: considerations for collecting and reporting race and ethnicity data in HCI

YIQUN T. CHEN, Stanford University, United States ANGELA D. R. SMITH, University of Texas at Austin, United States KATHARINA REINECKE, University of Washington, Seattle, United States ALEXANDRA TO, Northeastern University, United States

Engaging diverse participants in HCI research is critical for creating safe, inclusive, and equitable technology. However, there is a lack of guidelines on when, why, and how HCI researchers collect study participants' race and ethnicity. Our paper aims to take the first step toward such guidelines by providing a systematic review and discussion of the status quo of race and ethnicity data collection in HCI. Through an analysis of 2016–2021 CHI proceedings and a survey with 15 authors who published in these proceedings, we found that reporting race and ethnicity of participants is very rare (<3%) and that researchers are far from consensus. Drawing from multidisciplinary literature and our findings, we devise considerations for HCI researchers to decide why, when, and from whom to collect race and ethnicity data. For truly inclusive, equitable technologies, we encourage deliberate decisions rather than default omissions.

CCS Concepts: • Human-centered computing \rightarrow Empirical studies in HCI; • Social and professional topics \rightarrow Race and ethnicity.

Additional Key Words and Phrases: race, ethnicity, systematic literature review, HCI research, survey

ACM Reference Format:

Yiqun T. Chen, Angela D. R. Smith, Katharina Reinecke, and Alexandra To. 2023. Why, when, and from whom: considerations for collecting and reporting race and ethnicity data in HCI. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 24 pages. https://doi.org/10.1145/3544548.3581122

1 INTRODUCTION

As identities of study participants have proven to influence the uptake, experience, and benefits of technologies, the Human-Computer Interaction (HCI) community has made considerable efforts towards inclusive and diverse research practices over the past few years [1, 16, 25, 47, 104, 105, 110]. For instance, gender HCI has emerged as a mature subfield of HCI that focuses on how people of different genders interact with technology [109]. Likewise, HCI for development is a growing subfield that considers how designs and technologies interact with the under-resourced and economically disadvantaged communities [112, 116]. In this work, we focus on one such diversity dimension — race and ethnicity of participants in HCI — which has remained relatively under-explored in current research. In 2017, Schlesinger et al. [104] found that less than 0.1% of the papers in the CHI proceedings between 1981 and 2016 engaged meaningfully with race, compared to 0.2% and 0.6% for gender and socioeconomic class, respectively. Similar findings were reported on the basis of a quantitative content analysis of accepted papers in CHI 2006, 2011, and 2016 [47]. The authors highlighted

Manuscript submitted to ACM

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org. © 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

the importance of intersectionality (i.e., an identity framework that seeks to understand the complexity of multiple, overlapping, intersecting social identities [17, 28, 95]) when examining the composition of the participants in HCI research. In particular, they emphasized that HCI researchers should take an interest in understanding how various dimensions of participants' identities (e.g., race, gender, socioeconomic status) interact with each other and provided recommendations for deeper engagements with the resulting complex identities.

Other research has explored the issue of race, ethnicity, and bias in HCI research and the HCI community via a critical race theory lens [83, 107]. Ogbonnaya-Ogburu et al. [83] argued that racism is pervasive in social-technical systems and implored that HCI research should be attuned to the issue of race; they suggested that participation of under-represented minorities must be sought after in all research activities. Concerted efforts among HCI researchers also led to a workshop titled "engaging in race in HCI" in CHI 2020 that aimed to identify better practices for engaging with race and improving racial inclusiveness and equity in the broader HCI community [107]. This workshop allowed the community to begin assembling recommendations for like-minded researchers to discuss the role and implication of race in HCI, which ultimately led to a series of zines featuring the relationship among race, inclusiveness, and HCI research [91]. One of the zines in the series, in particular, urged HCI researchers, practitioners, and designers to consider race and the implication of race throughout their research process, and provided an infographic to aid the process of investigating and reporting race in HCI research. In addition, recent works in HCI have highlighted the importance of engaging with traditionally excluded groups in HCI such as Black women [43, 44, 93–95]. They collectively demonstrated how current technology marginalized Black people and how race, gender, and economic class influence the design choice of various pieces of technology. Finally, this paper builds directly on the "prequel" on race and ethnicity collection of participants in HCI research [96] by providing considerations on when, why, and from whom HCI researchers collect study participants' race and ethnicity.

In addition to HCI, other fields of research have explored the topic of race and ethnicity in their research as well. For instance, in medical sciences, the American Medical Association (AMA) Manual of Style states that "specifying the race or ethnicity of study participants can provide information about the *generalizability of the results of a specific study*"; therefore, it recommends reporting aggregate race and ethnicity for all study participants [33]. The American Psychological Association (APA) [3] has made similar suggestions for empirical studies in psychology. However, given the breadth of research interests and methodologies in HCI, we should neither thoughtlessly copy existing recommendations nor ignore established practices from other disciplines. In particular, HCI has the tradition of illustrating how groups of users interact with technology, which situates the field in a unique position to narrate and reflect on the lived experiences of racial and ethnic minorities through data [28].

When put together, the existing work calls for (i) a deeper understanding of the current practice of reporting race in HCI; and (ii) a guideline consisting of considerations and recommendations for when, why, and from whom to collect and report race in research. Therefore, our work makes strides toward this goal by answering the following research questions:

- RQ1: When are the study participants' race and ethnicity reported in HCI research?
- RQ2: What are some considerations that speak for and against collecting this information?
- RQ3: What are relevant considerations on how to collect, report, and use this information?

By answering the three research questions, primarily in the context of race and ethnicity in the United States, we make the following contributions to the literature:

- We provide an empirical analysis of the frequency of reporting CHI participants' race and ethnicity, showing that less than 3% of CHI papers in the proceedings from 2016 to 2021 have included such information. Moreover, nearly half of these papers were published in CHI 2021, suggesting a recent increase in research efforts on race and ethnicity data in HCI.
- Through a survey with authors who published in CHI and were affiliated with a U.S. institution during the time of publication, we summarize the motivations and considerations for and against reporting race and ethnicity in HCI research.
- We also synthesize existing discussions on racial and ethnic data in related fields to examine the potential benefits and (unintended) consequences of reporting racial and ethnic data in HCI research.
- Finally, we close with a set of considerations designed for HCI researchers to reason about racial and ethnic data collection and analysis; in particular, we call for careful interpretations of race and ethnicity data, which serves to narrate the lived experiences of racial and ethnic minorities when they interact with the technology under study. While our work does not claim or intend to provide a set of absolute and complete rules, we hope it will spark the discussion on the topic of collecting racial and ethnic information in the broader HCI community. After all, racial and ethnic data collection is a *start* rather than an end in combating systemic inequality, discrimination, and oppression in HCI, a journey requiring much caution and nuance from researchers to minimize harmful narratives and misrepresentations.

The authors humbly acknowledge that collecting and discussing race and ethnicity data is extremely complex and researchers' perspectives on this topic are greatly shaped by their own personal experience and research background. Collectively, we are U.S.-based HCI researchers with different research themes (large-scale cross-cultural online studies, narratives and participatory methods for more inclusive design, and statistical models), methodological focuses (quantitative, qualitative, design, and mixed-methods) and career stages (graduate students, early tenure-track assistant professors, and tenured associate professors). While our team has good coverage of HCI research methodologies, we are largely limited to specific social and cultural contexts of race and ethnicity in the U.S., where all authors currently work and live in. Thus, our work might have limited discussion on race and ethnicity in certain non-U.S. cultures and societies. We hope that our work, despite its U.S.-focus, could encourage broader conversations across research domains and cultural contexts (especially non-Western ones) in HCI.

2 UNDERSTANDING THE HISTORY OF RACIAL DATA COLLECTION

Our work is motivated by both the practice of racial data collection outside of HCI and the growing efforts to improve the collection of related demographic variables, such as gender and socioeconomic status [19, 49, 82, 102, 108] within HCI. In this section, we first briefly review racial categorization in the U.S., which provides the foundation for our quantitative analysis of CHI papers. Furthermore, we provide a selective overview of the practices for collecting and analyzing racial data in several research disciplines. We end the section by sketching how those approaches might inspire parallel efforts on racial data in HCI.

2.1 Racial and ethnic categorization in the U.S.

Racial categories have been included in every U.S. census since 1790, and the history of the U.S. census reveals the complexity of racial and ethnic data collection [88]. Firstly, prior to the 1960 census, an individual's race was determined by census enumerators (i.e., professionals who are hired to visit and survey residents to compile data for the U.S.

census), rather than through self-report. Moreover, categories used in the census changed almost every decade to reflect the politics and societal values at the time. For instance, "Native Hawaiian or Other Pacific Islander" was historically grouped with Asians and only became a new category in 2000. As another example, "Mexicans" were counted as a racial category in 1930, but that category has since disappeared. Since then, many of Mexican descents resort to the "other race" option and have been grouped under the Hispanic ethnicity only. The ability to identify as multi-racial (i.e. "mark one or more racial categories") was only won through extensive advocacy in 2000 [121]. Racial categorization has always been extremely political, and the miscategorization and undercount of people from racial minority groups have contributed to systemic oppression and exclusions [5]. In particular, these examples from the U.S. census illustrate that racial and ethnic data *collection* does *not* automatically lead to racial and ethnic *equality*. While the Census Bureau now collects racial and ethnic information to "make policy decisions for civil rights, to promote equal employment opportunities, and to assess racial disparities in health and environmental risks" [113], such data was in fact exploited by the government to *prosecute and oppress* racial minorities in the past [13].

U.S.-based researchers may also be familiar with the standards published by the U.S. Office of Management and Budget (OMB) in 1997, which mandates minimum standards for collecting and presenting data on race and ethnicity to support data analysis across different racial and ethnic groups. The OMB standards have two categories for ethnicity (Hispanic versus Non-Hispanic) and five categories for racial data at a minimum — White, Black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander. The OMB standards have been the guideline for collecting and presenting data on race and ethnicity for all federal reporting, including the decennial census and the mandates by certain U.S. grant funding agencies such as the National Institutes of Health (NIH) and the National Science Foundation (NSF) [79, 80]. Despite its general success in presenting data on race and ethnicity in federal programs, the OMB standards are far from a "gold standard", and performed especially poor on historically marginalized groups and/or multi-racial populations [62, 63].

From the history of racial categories in the U.S., we see that racial and ethnic data not only reflects the present social and political environment but also directly influences the future socialization and construction of the concept of race and ethnicity. Therefore, we conclude this section by emphasizing that the goal of collecting and reporting racial and ethnic data is not to arrive at a list of "perfect" racial and ethnic classifications, but more importantly, to leverage such data to reflect, acknowledge, and reject racist and oppressive power structures in current research.

2.2 Collection and analysis of racial data in research

In this section, we look at how the social sciences, medical sciences, and computer sciences have collected and analyzed racial data as comparative case studies.

Social science scholars have long acknowledged the role of race in shaping individuals' social status and everyday life experience [9, 39, 39, 69]. However, there is less consensus on whether the field should use racial classifications to assess the role and consequences of race. Some argue that collecting and reporting data on race and ethnicity would promote racial division and further the status quo of racial discrimination. In contrast, others take a "what we cannot measure, we cannot understand" approach and continue to report observed racial differences from profiling in law enforcement to disparity in healthcare systems [9, 14]. In 2003, the American Sociological Association (ASA) issued a statement in support of the continual collection and research of data on race [4]. Their reasoning is summarized as follows: (i) racial identities are central to the social organization and relationships and, therefore, the very core of social science research; (ii) taking a "colour-blind" approach and ignoring participants' race in research does not eliminate

the use of racial categories and racism in everyday life, as well as the resulting impact on societal outcomes; and (iii) understanding the role of race is central to challenging the existing systems of racial discrimination and stratification.

In contrast to the debates over racial data collection in social sciences, race and ethnicity of study participants are widely collected and used in healthcare databases to ascertain important group-level differences in healthcare outcomes in the U.S. [64, 86]. The basis of the observed race-associated differences in healthcare outcomes, however, remains under-explored, further illustrating that data collection itself does *not* automatically translate into racial and ethnic equality. One exception is Jones [59], who argued that as a social construct, race only serves as a very rough proxy for variables of interest such as social class and culture. Instead, race often appears predictive of healthcare outcomes because of the *racism* that has operated throughout U.S. history and to date. As an example, an analysis by Jones et al. [60] demonstrated that being classified by others as "White" is associated with better health status, *regardless of one's self-identification*. In view of the complex interpretation of "race", multiple threads of work have urged researchers in public health to take an interest in elucidating the underlying causes of the observed differences across race and ethnicity groups, e.g., by generating hypotheses about the basis and designing data collection and analysis plans to test the hypotheses. For instance, the difference in rates of estrogen-receptor-negative breast cancer between Black and White women in the U.S. is well-documented. Building on this observation, Krieger et al. [67] demonstrated that being born in the states that practiced Jim Crow laws (i.e., legal racial segregation) is associated with higher odds of cancer, thereby attributing the observed differences to racially discriminating laws.

In computer science, fair machine learning is one of many research areas that share similar interests in analyzing racial and ethnic data. With the growing role of predictive algorithms in critical fields such as credit reporting and employment assessment, a plethora of work has investigated the fairness of the algorithmic decisions across "sensitive attributes" such as different racial and ethnic groups [8, 11, 57]. Even though contextual discriminatory impacts of algorithms, such as the disparity between White versus Black Americans in the criminal justice system, are often cited as motivating examples for this line of research, researchers have largely treated racial or gender groups as an abstract feature of individuals. A few recent exceptions include Benthall and Haynes [8], Hanna et al. [42], and Hu and Kohler-Hausmann [50]. Benthall and Haynes [8] argued that people who are labelled as "Black" in the U.S. are subject to systemic differences through spatial, political, and social segregation. Therefore, the use of race categories such as "Black" might risk reifying racialized social inequality, if the observed differences are attributed to the race instead of to the underpinning systemic inequalities. Hence, Benthall and Haynes [8] proposed replacing racial categories with group labels learned through data to dynamically capture the underpinning inequalities, without furthering the status quo of disadvantaged racial groups. On the other hand, Hu and Kohler-Hausmann [50] investigated the social meaning of racial and gender membership. Their key insight is that prevailing analyses of socially salient categories such as race neglect the fact that many of the "effects" (e.g., career choices, family, and neighbourhood wealth) attributed to race are, in fact, its constitutive features. They implored for more care in conceptualizing and interpreting the "causal effects" of race. Finally, Hanna et al. [42] argued that most existing research ignores the "multi-dimensionality" of race and instead treated racial and ethnic categories as a fixed attribute. They urged researchers in the field to contextualize the meaning of race, and focus on the underlying mechanism that produces and reinforces racial inequality.

Within HCI, several authors have argued, through qualitative, mixed-methods, or quantitative methods, that there is a general lack of meaningful engagement with race and ethnicity [29, 47, 83, 91, 93, 104]. However, the current practice, as well as the motivations, for collecting and reporting study participants' race and ethnicity remains under-explored. In this work, we analyzed recent CHI proceedings to understand the existing practice. Furthermore, we also surveyed authors to identify the motivations and methods for collecting racial and ethnic data of their participants. Motivated by parallel efforts on the gender of study participants [82, 102, 108], we drew from related disciplines and outlined considerations for HCI researchers when it comes to racial data collection.

3 STUDY OF RACE AND ETHNICITY DATA COLLECTION IN HCI

To answer our research questions, we conducted a systematic literature analysis of the published papers in CHI proceedings from 2016 to 2021. We followed up with a survey of the authors whose papers reported race and ethnicity data of their participants in this period. The restriction on time is motivated by (i) the goal to understand the *current practice* on collecting racial and ethnic data of study participants, and (ii) the observation that older literature rarely reports this information [47].

3.1 Dataset curation

We started with a total of 3,910 research articles published in 2016-2021 CHI proceedings on ACM digital library and proceeded with a keyword-search method informed by prior work [47, 104]. We initially filtered the collection using the keywords "race" OR "racial" OR "ethnicity", which narrowed down the corpus to 663 articles. We then experimented with adding potentially defining keywords (race/racial, ethnicity/ethnic, White/Caucasian, African American/Black, Asian, Hispanic, Native American/American Indian, Pacific Islanders), and sampled the first 10 articles returned in each year to judge the quality of our search (the quality here is defined to be the number of search results that contained detailed race and ethnicity information of the study participants). The final keyword set that yielded the most relevant articles empirically is (i) "race" OR "racial" OR "ethnicity" AND (ii) "Hispanic" OR "Black American" OR "Asian" OR "Caucasian". After identifying this initial collection of 340 articles, we then checked for the racial and ethnic composition of study participants over all articles that detailed this information. Because of the focus on race and ethnicity in the U.S. context for this work, we further limit the corpus to the articles for which the authors had U.S. affiliations *or* the participants were recruited in the U.S. The dataset curation process is displayed in Figure 1.

For reference, we obtained racial and ethnic composition from two additional sources: (i) 2015–2019 demographics estimates of the U.S. collected by the United States Census Bureau [114]; and (ii) 2015–2019 demographics estimates of U.S.-based drug trials collected by the U.S. Food and Drug Administration (FDA) [12].

3.2 Dataset analysis

As stated in Section 2.1, categorizing race and ethnicity is extremely complex. In this paper, for simplicity, we adopted the following procedures to group the race and ethnicity categories across different studies. Firstly, since most of the surveyed CHI studies collected ethnicity and race using one single question for race and ethnicity, we make the assumption that the racial categories reported by studies in our final corpus refer to the corresponding Non-Hispanic subset (e.g., reported White participants in a paper refers to the non-Hispanic White participants). Second of all, we aggregated the participants into the following categories that roughly align with the OMB standards: White, Black or African American, Asian, and Others (including American Indian and Alaska Native, Mixed races, Native Hawaiian and Other Pacific Islander). This choice of analysis is largely driven by the existing racial categories in papers published in CHI, which certainly does *not* capture the full complexity of race and ethnicity of the study participants.

In addition to the aggregate analysis of studies in the final corpus, we will also report the following summary statistics: (i) the number of studies that collected ethnicity separately from race, and (ii) the racial and ethnic breakdown

of large (> the median number participants across all studies in the final corpus) studies and small-to-medium-scale studies (\leq the median number of participants across all studies in the final corpus).

3.3 Survey

While all papers in our final corpus reported their participants' race and ethnicity, many left the *reason and process* of collecting the race and ethnicity of their participants *implicit*. We, therefore, conducted an additional survey to find out why researchers collect racial and ethnic data and to learn about potential challenges they may have experienced. For each publication in the final curated corpus, we emailed the first and senior authors with the following list of open-ended questions on why and how they collected the race and ethnicity information of their participants.

- (1) Why did you decide to collect the racial and ethnic information of your participants?
- (2) How did you collect the racial and ethnic information of your participants? (i.e., what were the specific questions that you used?)
- (3) Did you use any resources to decide on whether and how to collect racial and ethnic information about your participants? (e.g., U.S. Census for U.S.-based studies, prior CHI publications, IRB recommendations)
- (4) If you happen to still have your original questionnaire and would be willing to share it with us, that would be greatly appreciated as well!

After excluding those authors with an inactive email address (e.g., due to a change of affiliation), we emailed a total of 106 authors (counting both first and last authors) whose papers were in our final corpus. Among those, 15 (14.2%) participated in our survey, all of whom were currently affiliated with a U.S. institution or corporation.

In addition to the survey responses, we note that some papers already highlighted the importance of considering race and ethnicity for the piece of technology under study. For instance, Passmore et al. [87] surveyed gamers from diverse racial and ethnic backgrounds and established "significant differences between players of color and White players on the perception of racial norms in gaming, effects of behavior, emotions, player satisfaction, engagement, and beliefs stemming from a lack of diversity." Moreover, they emphasized that the diverse recruitment amounted to "higher dissatisfaction [in diversity in digital games] than previous research."

4 RESULT

4.1 RQ1: When are the study participants' race and ethnicity reported in HCI research?

At the time of data collection, the ACM digital library includes 3,910 CHI papers published in 2016–2021 CHI proceedings. We analyzed 340 manuscripts that mentioned keywords related to race and ethnicity (see Figure 1 for details), of which only 93 provided descriptive statistics on the racial and ethnic breakdown of their study participants. In other words, our analysis showed that only 93 (2.4%) of 3,910 CHI papers included descriptive information on participants' race and ethnicity. This is likely an undercount given that the final corpus only included studies with the specified keywords.

Out of the 93 manuscripts, the median number of reported racial and ethnic groups is 4 (IQR: 3–5). Only a small number (17; 18.2%) of studies mentioned (or was inferred of) using two separate questions for race and ethnicity. The median number of participants in the studies in the final corpus is 28 (IQR: 18–187); the largest study reported the racial and ethnic breakdown of 2,041 participants [115], and the smallest study in our corpus had only six participants [40]. In addition, almost all (>90%) of the authors in our final corpus were affiliated with a U.S. institution at the time of writing. The number of manuscripts is also not evenly distributed across time, with more than 43 out of 93 (46%) published in CHI 2021, followed by 14 in 2017, 11 in 2019, 11 in 2018, 9 in 2020, and 5 in 2016.

The aggregated studies reported 19,684 participants in total, 12,627 (64.1%) of whom are (Non-Hispanic) White; 2,028 (10.3%) Black; 1,766 (8.9%) Hispanic; 1,327 (6.7%) Asian; and 1,939 (4.6%) Others (with 205 Mixed races and 98 American Indian or Alaska Natives). By contrast, according to the estimated demographic data by the U.S. census for 2015–2019, 60.7% of the population in the U.S. is (Non-Hispanic) White; 12.3% Black; 18% Hispanic; 5.5% Asian; and 3.5% Others (2.4% Mixed and 0.7% American Indian or Alaska Natives). Regarding the U.S.-based FDA drug trials during 2015 and 2019, (Non-Hispanic) White accounted for 64.5% of the participants, followed by 16% for Black and 15% for Hispanics. Asians and Other groups account for 2% and 3.5% of the trial participants, respectively.

Racial and ethnic compositions from the three different sources (CHI, U.S. Census, and the FDA drug trials) are displayed in Figure 2. We see that compared to the U.S. Census, CHI studies in our final corpus have slightly more Non-Hispanic Whites and slightly fewer Hispanics. In addition, Figure 2 suggests that participants in neither CHI studies nor FDA trials are representative of the aggregated U.S. demographics. However, we note that many CHI



* Keywords: (i) ``race" OR ``racial" OR ``ethnicity" AND (ii) ``Hispanic" OR ``African American" OR ``Asian" OR ``Caucasian". Articles not containing these keywords are excluded using the automatic search tools of the ACM digital library.

** Since an overwhelming majority of the articles that reported detailed racial and ethnic data recruited their participants from the U.S., we adopted this criterion to focus our discussion of race and ethnicity in the U.S.

Fig. 1. The flow of information through different phases of the corpus curation process as described in Section 3.1. We displayed the inclusion and exclusion criteria, as well as the final number of resulting publications of each stage. Here, articles screened for keywords (n = 340) are the articles that discussed or mentioned race and ethnicity regardless of whether detailed participant-level data are reported.

studies actively recruited a representative sample of their interest, which may or may not agree with the aggregated demographics of the U.S. For instance, Lopez et al. [74] was a Non-Hispanic-White-focused study, and Dosono and Semaan [24] specifically looked at the engagement and dynamics of the Asian American and Pacific Islander online communities. As a result, both studies will not resemble the U.S. demographics by design, rather than by omission.

We also looked at the longitudinal trend of compositions of reported racial and ethnic groups across the six years of CHI proceedings. Overall, the racial and ethnic compositions of study participants appear stable over the course of six years, with a more noticeable increase of Non-White participants from 2020 onwards. Figure 3 displays the racial and ethnic composition over the six years, stratified by the size of the study, where a study is classified as "large" if it has more than 28 (i.e., the median number of participants across all studies) participants and "small-to-medium" otherwise. We see that large-scale studies tend to have more White participants. This is partly due to the use of online platforms (e.g., Twitter or Mechanical Turk) for participant recruitment, which has been known to skew towards White samples [70, 117]. On the other hand, small-to-medium studies are more likely to target specific populations of interest (e.g., studying particular technology of interest in low-income neighbourhoods or among Black females; see, e.g., Ogbonnaya-Ogburu et al. [83] and Wheeler and Dillahunt [120]). As a result, small-to-medium studies might appear to have a larger proportion of non-White participants than larger studies.



Fig. 2. Racial and ethnic compositions of participants (in five groups) from three different sources: CHI proceedings (2016–2021, leftmost); demographics projection of the U.S. census (2015–2019, middle); participants of U.S.-based FDA drug trials (2015–2019, rightmost).

4.2 RQ2: What are some considerations that speak for and against collecting this information?

Responses from authors who participated in our open-ended survey are summarized in Table 1. Because the authors could cite multiple reasons in their open-ended responses, occurrences of the summarized categories in Table 1 may add up to over 15.

The most common reason ("Why" in Table 1) for collecting and reporting racial and ethnic information is *external validity*, that is, the degree to which the conclusion in one study would hold for other persons in other places and



Fig. 3. Racial and ethnic compositions of study participants from CHI proceedings between 2016 and 2021, by year and study size (large, >28 participants; small-to-medium, \leq 28 participants).

times [36]. For instance, one surveyed researcher noted in their response that "If my data is really only from a sample of white people, then I need to acknowledge that as a limitation of the study and ensure that my analysis is contextualized in that particular identity." A majority of the surveyed researchers also named prior work (including a priori hypothesis on the differences across racial and ethnicity groups) as a driving factor. As an example, one respondent noted that "We were specifically interested in experiences of intra-community marginalization, which includes systemic biases such as racism, and we wanted to ensure that our sample could capture such dynamics." In addition, the interplay between race, racism, and socioeconomic class in the U.S. motivates some researchers to "always collect these data" because when it comes to disparities as they affect technology, "race and income are so woefully correlated in this country [the U.S.] (and others) — it was important to collect racial/ethnic information." Only two out of 15 responses mentioned "external requirement" as the primary reason for collecting and reporting participants' race and ethnicity. Out of the two respondents, one stated that "I was in a Biomedical and Health Informatics program, and health studies often have people collect this data (perhaps tied to NIH funding/grant requirements)"; the other one mentioned that "our multi-year federal grant that funded this research had annual reporting requirements about, among other details, the demographics of our participants in each study". Both responses speak to the possibility of leveraging practices and training from related disciplines, such as biomedical sciences, to improve the collection and report standards of race in HCI.

Regarding the method of collection ("How" in Table 1), all of the surveyed researchers administrated a questionnaire, but the details of the administration varied: The same number of authors used two separate questions (seven out of 15) versus one single question (seven out of 15) when obtaining the race and ethnicity information of their study participants. This is in contrary to the observation made from our systematic literature analysis, where less than 20% of the studies used two separate questions. Several researchers also used a combination of categorized and open-ended responses, where the study participants can provide their own race and ethnicity.

	Occurrence
Why	
External validity	8
Targeted studies	4
Motivated by prior work	8
External requirement	2
Motivate future studies	4
How	
Separate questions for race and ethnicity	7
One question for race and ethnicity	5
Open-ended responses	4
Reference	
U.S. Census	8
Sociology research	1
A priori population of interest	3
Pilot study and prior work	4

Table 1. A summary of surveyed authors' responses. Note that multiple reasons and sources are allowed. Therefore, occurrences in each individual category could add up to more than 15. The categories are determined through qualitative coding of the responses.

4.3 RQ3: What are relevant considerations on how to collect, report, and use this information?

All responses mentioned existing resources that informed their data collection process (see the "Reference" column in Table 1). Among those responses, eight out of 15 cited the U.S. census as the primary reference for designing the categories in the questionnaires. However, researchers are aware of the limitations of U.S. census as a reference, "*I found the categories somewhere* — *either NIH or census perhaps but unfortunately not sure. I know that I tried to find existing categories that the government recommended as I thought these would be the absolute best practice/best way of going about this and then later learned these may not tie best to how people identify themselves.*"

Another response mentioned categories informed by the social sciences, partly due to their expertise in race and ethnicity research — "we reviewed a number of articles from top sociology journals in the years immediately prior to the data being collected (e.g., ASR [American Sociological Review], AJS [American Journal of Sociology]), as sociologists tend to think more about these issues, and also use representative survey data, than communication scholars." In addition to the categories informed by the U.S. census, a few studies also identified a target population of interest via prior work, e.g., one respondent noted that "we looked at previous research in this area and consulted with our community partners who collect this information as part of their data collection and capacity building activities."

5 CONSIDERATIONS

In the following, we discuss some considerations for collecting race and ethnicity data, synthesized from our multidisciplinary literature review *and* survey data.

5.1 Why: Whether to collect racial and ethnic data

Answering the question of whether to collect racial and ethnic data of study participants requires substantial care. On the one hand, researchers in HCI and many related scientific disciplines are trained and required to justify the collection and planned analysis of demographic variables such as race. Documented reasons for not collecting this

data may be privacy concerns for the study participants [6], or preventing survey disengagement and fatigue [38, 51]. Moreover, some fear that collecting and reporting on the findings in different racial groups, when presented without a thorough and nuanced discussion, could reify the racial inequality and stereotype, and even delve into "scientific racism" (i.e., the pseudoscientific belief that empirical evidence justifies racial discrimination) [4]. When put together, these concerns result in recommendations such as "Researchers should not collect more information from participants than is needed to answer their research questions" by the Institutional Review Board (IRB), an institutional organization that approves, monitors, and reviews human-subject studies in U.S. academic research institutions [10]. By contrast, demographics including gender, age, and whether participants are in vulnerable groups (prisoners, minors) are routinely requested in human-subject studies applications to IRB. In practice, researchers often need to cite established differences in racial groups to justify their decisions. On the other hand, one cannot hope to find a documented "established difference" if no study is devoted to documenting and understanding potential differences of the piece of technology across racial groups. This "what we do not measure, we do not understand" argument is especially relevant, given that race has been understudied, and the historic impact of race (and racism) in the U.S. and many countries have been made invisible. For example, in the U.S., the NIH started mandating the recruiting and reporting of racial and ethnic minorities in all clinical trials after observing significant differences between racial groups across a wide range of health care outcomes [34, 65, 78]. While federal agencies in the U.S. (such as the NIH and the Census Bureau) have generally leaned towards collecting race and ethnicity of participants, the legality and regulations of the collection and analysis of racial and ethnic data differ greatly around the world. For instance, in the European Union (EU), the General Data Protection Regulation (GDPR) [18] mandates that "processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs ... shall be prohibited [in general]." Despite some listed exceptions, e.g., "archiving purposes in the public interest, scientific or historical research purposes or statistical purposes", the legal and logistical challenges led to another important consideration for participants and researchers based in the EU (and other regions with similar laws and regulations).

Given the nuanced nature of this topic, we encourage HCI researchers to have an active discussion during the study design phase within the team. Some points of discussion (as framed through quotations from our survey respondents and findings from our review) may include:

- (1) Observed differences in the outcome of interest across racial groups:
 - e.g., Yardi and Bruckman [123] demonstrated that low-income African-American families share digital devices more often than their middle-class White-American counterparts, motivating future work (e.g., Garg [35], Pina et al. [89]) to consider the effects of race and socioeconomic status on technology usage.
 - e.g., One participant studied "the use of [ICTs] within and among online communities engaging in identity work". Since their outcome of interest is interwoven with race and ethnicity, "collecting the racial and ethnic information of participants contextualizes the results within the perspective of the moderator's background".
- (2) Potential consequences and implications of the study:
 - e.g., Collecting and reporting racial and ethnic information could also be useful for *"future studies that may want to compare their results with ours "*. While their study mainly examined the effect of gender, they still *"wanted to enable any future studies examining race to be able to compare their context with ours."*
- (3) Race and ethnicity of study participants is an important dimension of the external validity of a study:
 - e.g., Some researchers "always collect these data. As a person trained as a quantitative researcher, this seems like basic, key demographic information that one should have on-hand in case it is relevant to the RQs."

- e.g., another respondent reflected, "the 2016 paper had a sample that was over 90% white, which is definitely not diverse or representative of the trans population [their population of interest]." Therefore, in their follow-up work in 2020, they made sure to "made sure to prioritize recruiting a diverse group on many dimensions, with a particular focus on race/ethnicity".
- (4) Concerns for privacy and race "determinism":
 - e.g., Widespread collection of individual race and ethnicity data may spark privacy concerns. The research team should communicate the intended use of the collected data, and work with the community to understand whether the benefits outweigh the privacy and confidentiality concerns.
 - e.g., Presenting data and findings with racial and ethnic identities may also risk perpetuating theories of racial
 or genetic determinism (i.e., the belief that genetics or phenotype exclusively account for human behaviour
 and ability). This is echoed by one of our survey respondents: "I think it's important to choose to include these
 variables [race and gender variables] in the analyses very carefully. I think race and gender variables especially
 are sometimes overused because when differences arise they can give way to post-hoc explanations that reinforce
 stereotypes. That is, the use of these variables should be thoughtful and intentional."

Of course, given the nuanced nature of this question, we are *not* arguing to simply lower the barriers to collecting racial and ethnic data of participants for all HCI studies — as noted above, concerns for privacy and misinterpretation of the results should not be taken lightly. Instead, we advocate that whether or not to collect racial and ethnic data of participants should be *a deliberate* rather than *a default* decision.

5.2 How to collect race and ethnicity data

Our findings show that researchers who have committed to collecting race and ethnicity of their participants desire a standardized method of collection and seek out templates and best practices. Because of our restriction to U.S.-based participants and research, the vast majority of researchers in our study used the U.S. census as their references when designing their surveys and questionnaires (see Table 1). A considerable subset of surveyed researchers also used the modified format to ask one combined question about race and ethnicity, rather than two separate questions. While this may be an effective starting place, the U.S. census is only conducted every 10 years, and its categories often represent the political values at the time rather than the best research practice [5, 69]. Therefore, there is a need for best practice that is research-driven and takes into account how participants actually want to be identified by race. In this section, we offer some specific, practical advice on how to collect participants' race and ethnicity using surveys.

As an example, consider the collection of race and ethnicity in the U.S. census — census forms now have two separate questions about race (i.e., what is this person's race) and ethnicity (i.e., whether this person is of Hispanic, Latino, or Spanish origin). This separation was largely motivated by the growing diversity within the Hispanic-American population. However, in the 2010 Census, 37% of surveyed Hispanic or Latinx chose not to identify with any of the provided race categories [52], which reveals the gap between the provided categories in the status quo and how people identify themselves.

While there is certainly no one-size-fits-all solution to collecting race and ethnicity of study participants, we hope the following pointers could be of help to researchers during data collection:

(1) When the race and ethnicity information is collected via multiple choice questions, the research team should consider allowing the option of *identifying with more than one race and ethnicity*, and including an *open-ended option* for participants to self-describe their race and ethnicity [20, 98].

- (2) If researchers want to provide racial and ethnic categories in a questionnaire (e.g., in a large-scale online study), the U.S. census is a good starting point for U.S.-based studies (e.g., in the U.S. context, consider asking separate questions for race *and* ethnicity). However, depending on the nature of the study, categories used in the census (e.g., Asian Americans and Pacific Islanders) do not necessarily capture the underlying diversity of the group, and more granular choices might need to be included to reflect and communicate participants' identities [48, 53, 62]. As one of our survey participants put it, "highlighting the inner diversity of the sample also communicates that the AAPI [Asian Americans and Pacific Islanders] umbrella should not be construed as a monolith (e.g., the socioeconomic experiences of East Asians vary from Southeast Asians)."
- (3) When there are limited resources on racial and ethnic categories of participants for a study, consider asking participants to self-describe and then cluster the responses afterwards, which could serve as a starting point for future studies engaging with similar populations. In resource-limited settings, researchers could employ similar strategies in a smaller pilot study with a subset of participants. If resources permit, we also encourage researchers to solicit feedback from participants on whether the proposed categories capture their identified race and ethnicity.
- (4) In addition to race and ethnicity, consider alternative data that captures a particular dimension of racial differences and could address the research questions of interest more directly: languages spoken at home, household disposable income, access to quality healthcare (or the lack thereof), etc. For instance, Hanna et al. [42] and Roth [99] have outlined some of these particular dimensions or "proxies".

5.3 From whom to collect race and ethnicity data

Even equipped with perfect survey questions, researchers are likely to run into a series of additional practical challenges — Who should be recruited for the study? Should we always strive for racially-representative samples? How to recruit and retain racial and ethnic minorities? While answers to these questions will be context- and study-specific, we briefly summarize some common motivations and barriers for recruiting racially- and ethnically-diverse or homogeneous participants in this section.

Recall that many of our surveyed researchers cited the external validity of their studies as a primary motivation for collecting participants' race and ethnicity (see Table 1). Indeed, HCI researchers routinely draw inferences about populations by extrapolating findings based on data from small samples of people. Recent years have seen a surge of work on how studies in HCI, along with psychology, might have relied too much on samples from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) samples [7, 45, 73, 90] and among college students [41, 46, 90].

Therefore, we encourage researchers to assess the external validity, as it relates to participants' race and ethnicity, of their specific studies. For instance, Obiorah et al. [81] proposed and evaluated a novel interactive multi-person tabletop exhibit in museums, and the race and ethnicity of participants they engaged *"is reflective of typical museum visitors" demographics in the U.S.*", which speaks to the applicability of their findings on the populations they are interested in. By contrast, Hamidi et al. [40] acknowledged the lack of diversity in their samples as a limitation and stated that *"Our sample includes diverse age and gender perspectives but lacks diversity of race or ethnic identity [with respect to the U.S. demographics]."*

Of course, researchers do not always need to recruit participants to match national (or regional) race and ethnicity composition — external validity should be assessed with respect to the population from whom we are drawing inferences. For instance, F. Maestre et al. [30] recruited a majority of White participants, in keeping with the demographics of the rural region of the U.S. Midwest. Other notable examples include studies with a focus on selected racial and ethnic

groups: Lee and Rich [68] recruited primarily Black American participants because of the "substantial history and contemporary issues with medical racism towards the Black community". Similarly, Dosono and Semaan [24] were interested in the dynamics of online Asian American and Pacific Islanders communities. In both cases, a lack of racial diversity does not necessarily pose threat to the external validity of the results.

We also want to acknowledge that relying on convenient (and, often WEIRD and non-racially-diverse) samples has its practical advantages, because recruiting and retaining racial and ethnic minorities presents logistical, financial, and sometimes even legal barriers [6, 15, 18, 85, 90]. Examining the effectiveness of sampling and recruitment strategies tailored to specific ethnic minority groups (e.g., F. Maestre et al. [30], Sadler et al. [100]) in HCI is an important thread of future work. In practice, the following questions summarized and adapted from IRB forms at various universities might be an effective starting point for researchers who want to reflect on their study design and sample collection [54–56, 97]:

- (1) What is the estimated male-to-female ratio in the study? Does this reflect the distribution of the local population?
- (2) Is there any target population in terms of gender, race, ethnicity, sexual orientation, literacy level, health status, and economic class?
- (3) Is there any vulnerable population in your study who might be disproportionally affected by the research (e.g., indigenous people, minors, students)?
- (4) Will your study participants be representative of the demographics in the study region? Include an estimate of the percentages that will be from minority groups, or the lack thereof.
- (5) In addition, identify any racial, ethnic, or gender groups that will be specifically excluded from this research study. Consider providing a compelling justification for such exclusion (in addition to convenience samples).
- (6) If the research involves the collection of sensitive, potentially identifiable information (e.g., sexual orientation, race, gender, age, and a combination of the aforementioned factors), consider describing what information will be obtained and included in the final research output, and how permission will be sought.

While we focused on practical suggestions for collecting race in simplified settings such as surveys in this section, there is an array of important and exciting work dedicated to collecting complex, qualitative racial data through interviews, focus groups, and case studies [29, 83, 84, 93, 94, 120]. We encourage researchers to explore listed papers and references therein to engage with qualitative racial data collection.

5.4 What to report on racial and ethnic data

HCI researchers might also find themselves analyzing race and ethnicity data of their participants, such as incorporating participants' race in a regression model. While a well-conducted analysis will enrich our knowledge of the technology under study, such analyses, if done poorly, might give way to *post hoc* explanations that reinforce racial stereotypes and inequality [27, 32]. Below, we outline a few considerations for reporting and communicating the findings based on the race and ethnicity of study participants. These considerations are largely inspired and adapted from existing discussions in medical sciences [33, 59, 64], fair machine learning [6, 8, 50], HCI [29, 43, 93, 95, 102], and our own experience as researchers in the field.

Firstly, we encourage researchers to acknowledge that race is a social construct and proxy, rather than an objective measure of underlying traits [8, 50, 59] when presenting observed differences in racial groups. In particular, we caution *against* languages that suggest the differences in the measured outcomes can be attributed to participants' race, a slippery slope to furthering the racial disparity and stereotypes. Instead, authors should provide qualifying statements when presenting their results across racial categories. As an example, in CHI 2021, K. Chua and Mazmanian [61] stated

that "We recognize that the terms 'Asian,' 'White,' and 'Hispanic' are broad and homogenize the experiences of people from various races, ethnicities, and national backgrounds. We chose to use these terms because they are emic terms used by the vast majority of our participants in describing themselves."

Second of all, we urge researchers to investigate the potential basis of the observed differences across racial and ethnic groups to the extent feasible, e.g., conducting a follow-up study or drawing from existing literature. For instance, when Dillahunt et al. [23] found that participants of different races had different levels of engagement and outcomes with online employment resources, they drew from the existing literature on Black-White income inequality [37, 118] and hypothesized that the observed differences across race could be attributed to the design of online resources websites with the "Us versus Them" thinking that marginalized the racial minorities.

Thirdly, discussions about the potential impact of the study should be sought after. This effort could take many forms: for example, the research artifact could include a paragraph on the societal impact of the research. Such practice would be widely applicable to research artifacts that propose new systems or software, paralleling efforts in the artificial intelligence community [2]. As an example, if a paper proposes a new piece of virtual reality technology, and researchers observe that the accuracies of the technology differ across racial groups, they should discuss the potential impact when the technology is deployed at a large scale. Similarly, for research outputs that incorporate machine learning models as part of their systems, potential disparate impacts on minority groups, as well as a toolkit for mitigating such ramifications in fair machine learning, could be discussed [6, 8].

Additionally, we want to call attention to the default practice of treating racial groups as categorical variables with "White" as the reference group (e.g., by classifying participants as "White" versus "Non-White") in quantitative analyses in HCI. First of all, while easy to the implement, the grouping approach treats race as a non-overlapping attribute and ignores the heterogeneity within each racial and ethnic category [26, 42]. This could contribute to further reification of the groups under study, especially when the groups exhibit large internal differentiation such as the "Asian Americans and Pacific Islanders" category in the U.S. census [26]. As an alternative, we encourage quantitative analysis to account for the "multi-dimensionality" of race and carefully examine which dimension or representation is the most appropriate [42, 99]: for instance, researchers might adopt phenotypes such as skin type in investigating computer vision applications Buolamwini and Gebru [11]; or geography-based information such as those proposed in Benthall and Haynes [8] when looking at access to public resources. Furthermore, even when census-like racial categories are suitable, researchers should reflect whether they are defaulting to a culturally-dominant group (e.g., male, White, high income) as the reference could subtly establish and reify the notion that culturally-dominant groups are the most "normal" and "interesting", thereby creating a nested system of importance over time [27, 32, 58]. Moreover, using the culturally dominant group as the reference does not always lead to the most interpretable results - for instance, if we are really interested in how different racial and ethnic groups compare against the average in the study population, coding the differences against a specific reference group is not the most informative approach (see more technical details and alternative modelling strategies in Dupree and Kraus [27] and references therein).

Last but not least, drawing from the unjust history of the U.S. census, as well as the thread of excellent work exposing the oppressive nature of current technology design in HCI [28, 43, 44, 94, 95, 103], we would like to emphasize that data collection and reporting is *not* sufficient in itself — researchers should always (i) consider narrating lived experiences of the racial and ethnic minorities; (ii) explore how oppression from race, gender, and class operates under the current technology design; and (iii) actively resist the current design and technology that perpetuates racism, whenever applicable.

5.5 Moving beyond the U.S.

In this paper, we primarily engage with race and ethnicity in the U.S. context. However, focusing on the North American experience does not fully embrace the issues of race and racism globally. Therefore, in this section, we describe how our analysis and reflections could shed light on how HCI researchers can consider why, when, and how to collect data about race and ethnicity globally:

- Why and from whom: our considerations for why and from to collect race and ethnicity in Sections 5.1 and 5.3 generalize well to non-U.S. contexts. For instance, socioeconomic and health disparities, and their subsequent impact on HCI research, remain persistent across different parts of the world [31, 111, 119]. In addition, the external validity of a study is increasingly important in a global context, as the field of HCI continues to expand beyond its focus on WEIRD samples and discovers more country-to-country, region-to-region differences [72, 73].
- How: as shown in Table 1, most U.S.-based researchers turn to the U.S. census for categorizing their study participants' racial identities. While other countries and regions might not have such clear-cut categories, our recommendations in Section 5.2 still point to a general path forward: researchers could start with existing resources such as the national or region census [106] and scholarly work in related disciplines [22, 111, 119, 124]; pilot studies could be especially helpful in a historically under-explored community in HCI.
- What: in addition to the considerations in Section 5.4, we want to highlight that collecting and reporting ethnic identities of participants around the world is merely a start to more equitable, global, and diverse HCI research. Especially with a sample of non-U.S. participants, researchers need to pay particular attention so that the research outputs and insights do not remain WEIRD-focused [77, 122].

6 DISCUSSION

The primary goal of our work is to understand the race and ethnicity data in HCI from the following aspects: (i) *when* are the HCI study participants' race and ethnicity collected (RQ1); (ii) *why* are race and ethnicity collected (RQ2); and (iii) *how* is the collection administrated, as well as curating a list of considerations and recommendations for collecting and reporting race and ethnicity in the future (RQ3). Our analysis revealed that for studies published in CHI between 2016 and 2021, less than 3% included detailed race and ethnicity information about their study participants. Among those studies that are based in the United States, about 64% of total participants identified as Non-Hispanic White. By contrast, 9% and 10% identified as Hispanic and Non-Hispanic Black, respectively. Regarding "why", our participants cited the themes and historical contexts of their research as the primary reason. Other motivating factors to collect this information are to increase a study's external validity, achieve a more representative sample, or allow future research. Lastly, regarding how the racial data were collected, we found that the U.S. census, as well as identified groups in prior work, were some of the most commonly cited sources.

Motivated by the literature review and results of our survey, we also compiled a list of considerations – corresponding to the three RQs – for when to include study participants' race and ethnicity (RQ2), and if so, who to include (RQ1) and how to collect this information (RQ3). Our considerations expanded on the pioneering recommendations by Ogbonnaya-Ogburu et al. [83] and Race in HCI Collective et al. [91], and are inspired by efforts in data collection and report of gender, which led to a crowd-sourced, pragmatic working document [102] serving to approach gender inclusively in HCI research. In particular, we advocate for *deliberate decisions* when determining when and from whom to collect race and ethnicity information. Furthermore, while the U.S. census, as well as its foreign counterpart, might serve as a good starting point for collecting this information, HCI researchers should stay attuned to the complexity of race and ethnicity

 more granular categories, obtained through related works or pilot studies, could lead to historically-neglected insights and minority-empowering research [106].

We also want to emphasize the intention of this paper: instead of a panacea for engaging with racial data, or the lack thereof, in HCI, this paper is meant to encourage more conversations on collecting and reporting race and ethnicity among HCI researchers, practitioners, and users including the participants in HCI research. Even within the team of authors, we have divergent opinions on some of the considerations, partly due to our diverse research themes, methodological focuses, and first-hand experience with collecting and reporting racial and ethnic information in HCI research. When it comes to when to collect race and ethnicity from study participants, one author who has extensive expertise in engaging with race in HCI immediately took a strong stance and argued that race and ethnicity of study participants should always be collected for a complete contextualization and narrative of the study. However, another author on the team challenged this stance by citing their own experience in collecting race and ethnicity for large-scale, multi-national online studies — a lack of community-driven and research-driven guidance on collecting racial and ethnic data, especially in non-Western contexts, poses a substantial logistical and methodological challenge in their quantitative work. In addition, the team's collective experiences with IRBs at different institutions also revealed a lack of standardized institutional guidance on collecting this information.

While the team of authors cannot and, certainly, does not claim to represent all research areas in HCI and identities experienced by HCI researchers, we see the tension among us as an indicator of the potential tensions in other research teams and the HCI community, highlighting that race and ethnicity in HCI is a nuanced and highly personal topic that needs to be handled with care. We believe that tension and discussion on collecting and reporting racial and ethnic data should be welcomed: the *tensions* among researchers and practitioners today could give rise to safe, inclusive, and equitable *resolutions* for the HCI community tomorrow.

7 LIMITATIONS

Our work is subject to several limitations. One limitation is the scope of the discussion of race and ethnicity: the paper and existing work surveyed herein are based on the racial and ethnic context of the United States. In part, this is due to the vast collection of existing research on race and ethnicity in the U.S. However, given the high research output of U.S.-based HCI researchers [71], we hope that our work will serve as a proof-of-concept piece for future conversations about race and ethnicity in a global context.

In terms of research methodology, our sampling could be subject to selection bias — while CHI covers a broad range of research topics, published papers in CHI proceedings are a small subset of the broader HCI research outputs. For instance, we might expect conference proceedings with a more focused theme (e.g., user experience conferences such as NN/g) and conferences with a more explicit international, non-Western focus (e.g., CLIHC and Asian CHI) to have different approaches to collecting and reporting race and ethnicity of study participants in HCI research. In addition, even within CHI proceedings, limiting publications to 2016–2021 also potentially confounds our findings with the longitudinal trend of research on race and ethnicity in HCI. For instance, more recent research outputs might have more discourse on race and ethnicity [47], and the year 2021 consists of the largest amount of papers in our final corpus. Therefore, if we repeat our methods on the CHI proceedings from 2022, there is likely going to be a substantial amount of papers we would include. Our approach to corpus curation can also lead to an undercount: there is an array of excellent work in CHI on race and ethnicity in HCI without explicitly recruiting participants and collecting their race and ethnicity. These papers are likely to be omitted in our curation process. Furthermore, as the primary intention of this work was to start a discussion on race and ethnicity data collection, we did not prioritize an exhaustive, iterative

refining of our corpus. Therefore, putting all these factors together, we postulate that the reported results on the final corpus of papers are likely an underestimate of the CHI publications that collected and reported participants' race and ethnic information.

Regarding the survey results, the 15 researchers who provided prompt responses to our inquiries might not be a representative sample of HCI researchers. For instance, 47% of the researchers who participated in our survey used separate questions to collect race and ethnicity, as opposed to less than 20% of the researchers in the entire final corpus. In addition, the collection and report of racial and ethnic data without a strong justification may be discouraged by regulatory agencies. Therefore, the information curated from published studies does not necessarily represent the initial study design or intentions of the researchers, but is likely a combination of the research agendas and constraints imposed by external resources and regulations. As one surveyed researcher noted in their response, "*We ran into challenges and limitations in mapping the participant responses to our grant reporting requirements. Therefore, we ended up collecting the racial/ethnic information based on US Census Bureau categories.*"

8 CONCLUSION AND FUTURE WORK

As HCI continues to engage with a racially- and ethnically-diverse population of users, understanding the current practice of collecting the race and ethnicity of participants in HCI research takes on high importance. Through a systematic review of published CHI papers and follow-up surveys with selected authors, we found that less than 3% of CHI papers from 2016 to 2021 collected and reported their participants' race and ethnicity. Among those authors who did collect this information, the primary motivations include (i) strengthening the external validity of the study, and (ii) addressing the established disparities in the uptake and use of technologies between different racial groups. Most surveyed authors mentioned the U.S. census as their reference for designing the questionnaires for collecting participants' race and ethnicity.

Our findings reveal several important directions of future work. Firstly, CHI is a global community and reporting on the ethnicity of participants outside of the U.S. has been steadily increasing (for instance, David Bowman et al. [21], Koushki et al. [66], Randhawa et al. [92]). Extending our studies to a more global context will champion the call for inclusiveness and representation of non-Western samples in the HCI research community. In addition, even in the U.S. context, the nuance of the racial groups is not necessarily captured by the established categories used in the U.S. census. For instance, although Middle Eastern and North African Americans are classified as White in the U.S. census, a sizable subset of the community believes that they are not treated or perceived as White, and that such classification might even perpetuate further harm [75, 76, 101]. Moreover, depending on the nature of the study, categories used in the U.S. census do not necessarily capture the underlying diversity within the group, and researchers have called for more granular categories might need to be included to reflect and communicate participants' identities [48, 53, 62].

Another avenue of future research is to investigate the challenges encountered in decisions around race and ethnicity data collection and analysis, especially among the researchers who decided *not* to collect and report such data. For instance, a systematic summary of the primary barriers (e.g., privacy and legal concerns, lack of systematic categories for large-scale international studies) could inform future efforts on providing resources and designing tools to overcome these barriers.

In this paper, we also outlined a few recommendations which serve to further the conversations on whether this data should be collected, and in what circumstances. Critically, we are not proposing that every HCI study should simply collect race and ethnicity of its participants. Rather, we want to highlight the importance of a deeper and broader consideration of racial and ethnic data collection and analysis in HCI, and certainly within the research team — as long

as racial and ethnic categories continue to govern social, political, and cultural interactions, collecting and analyzing racial and ethnic data fits squarely within the agenda of HCI.

REFERENCES

- Julio Abascal and Colette Nicolle. 2005. Moving towards inclusive design guidelines for socially and ethically aware HCI. Interacting with Computers 17, 5 (Sept. 2005), 484–505.
- [2] Grace Abuhamad and Claudel Rheault. 2020. Like a Researcher Stating Broader Impact For the Very First Time. arXiv (Nov. 2020). arXiv:2011.13032 [cs.CY]
- [3] American Psychological Association. 2019. Publication Manual of the American Psychological Association: 7th Edition, 2020 Copyright (7 ed.). American Psychological Association.
- [4] American Sociological Association. 2017. The Importance of Collecting Data and Doing Social Science Research on Race. https://www.asanet.org/ importance-collecting-data-and-doing-social-science-research-race. Accessed: 2021-8-5.
- [5] Margo Anderson and Stephen E Fienberg. 2000. Race and ethnicity and the controversy over the US Census. Current Sociology 48, 3 (2000), 87-110.
- [6] Mckane Andrus, Elena Spitzer, Jeffrey Brown, and Alice Xiang. 2021. What We Can't Measure, We Can't Understand: Challenges to Demographic Data Procurement in the Pursuit of Fairness. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 249–260.
- [7] Jeffrey J Arnett. 2008. The neglected 95%: why American psychology needs to become less American. American Psychologist 63, 7 (Oct. 2008), 602–614.
- [8] Sebastian Benthall and Bruce D Haynes. 2019. Racial categories in machine learning. In Proceedings of the Conference on Fairness, Accountability, and Transparency (Atlanta, GA, USA) (FAT* '19). Association for Computing Machinery, New York, NY, USA, 289–298.
- [9] Jack M Bloom. 2019. Class, Race, and the Civil Rights Movement, Second Edition. Indiana University Press.
- [10] Institutional Review Board. 2020. Racial Equity Considerations and the Institutional Review Board Child Trends. https://www.childtrends.org/ publications/racial-equity-considerations-and-the-institutional-review-board. Accessed: 2021-8-27.
- [11] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability and transparency. PMLR, 77–91.
- [12] Center for Drug Evaluation and Research. 2019. Drug Trials Snapshots. Accessed: 2021-3-8.
- [13] Robert Chang and Lorraine Bannai. 2019. Brief of Norman Y. Mineta, the Sakamoto Sisters, the Council on American-Islamic Relations (National and New York, Inc.), and the Fred T. Korematsu Center for Law and Equality as Amici Curiae in Support of Respondents. (2019).
- [14] Thandeka K Chapman. 2013. You can't erase race! Using CRT to explain the presence of race and racism in majority white suburban schools. Discourse: Studies in the Cultural Politics of Education 34, 4 (Oct. 2013), 611–627.
- [15] Meghan Coakley, Emmanuel Olutayo Fadiran, L Jo Parrish, Rachel A Griffith, Eleanor Weiss, and Christine Carter. 2012. Dialogues on diversifying clinical trials: successful strategies for engaging women and minorities in clinical trials. *Journal of Women's Health* 21, 7 (July 2012), 713–716.
- [16] Derrick L Cogburn. 2003. HCI in the so-called developing world: what's in it for everyone. Interactions 10, 2 (March 2003), 80-87.
- [17] Patricia Hill Collins and Sirma Bilge. 2020. Intersectionality. John Wiley & Sons.
- [18] European Commission. 2016. General Data Protection Regulation (GDPR) Official Legal Text. Accessed: 2021-9-8.
- [19] David I Conway, Alex D McMahon, Denise Brown, and Alastair H Leyland. 2021. Measuring socioeconomic status and inequalities. In Reducing social inequalities in cancer: evidence and priorities for research.
- [20] Paul R Croll and Joseph Gerteis. 2019. Race as an Open Field: Exploring Identity beyond Fixed Choices. Sociology of Race and Ethnicity 5, 1 (Jan. 2019), 55–69.
- [21] Nicholas David Bowman, Jihhsuan Tammy Lin, and Chieh Wu. 2021. A Chinese-Language Validation of the Video Game Demand Scale (VGDS-C): Measuring the Cognitive, Emotional, Physical, and Social Demands of Video Games. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–10.
- [22] Francis M Deng. 1997. Ethnicity: An African Predicament. The Brookings review 15, 3 (1997), 28-31.
- [23] Tawanna R Dillahunt, Aarti Israni, Alex Jiahong Lu, Mingzhi Cai, and Joey Chiao-Yin Hsiao. 2021. Examining the Use of Online Platforms for Employment: A Survey of U.S. Job Seekers. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–23.
- [24] Bryan Dosono and Bryan Semaan. 2019. Moderation Practices as Emotional Labor in Sustaining Online Communities: The Case of AAPI Identity Work on Reddit. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–13.
- [25] Susan M Dray, David A Siegel, and Paula Kotzé. 2003. Indra's Net: HCI in the developing world. Interactions 10, 2 (March 2003), 28-37.
- [26] Lucas G Drouhot and Filiz Garip. 2021. What's behind a racial category? Uncovering heterogeneity among Asian Americans through a data-driven typology. RSF: The Russell Sage Foundation Journal of the Social Sciences 7, 2 (2021), 22–45.
- [27] Cydney H Dupree and Michael W Kraus. 2022. Psychological Science Is Not Race Neutral. Perspectives on Psychological Science 17, 1 (Jan. 2022), 270–275.

- [28] Sheena Erete, Aarti Israni, and Tawanna Dillahunt. 2018. An intersectional approach to designing in the margins. Interactions 25, 3 (April 2018), 66–69.
- [29] Sheena Erete, Yolanda A Rankin, and Jakita O Thomas. 2021. I Can't Breathe: Reflections from Black Women in CSCW and HCI. Proc. ACM Hum.-Comput. Interact. 4, CSCW3 (Jan. 2021), 1–23.
- [30] Juan F. Maestre, Tawanna Dillahunt, Alec Andrew Theisz, Megan Furness, Vaishnav Kameswaran, Tiffany Veinot, and Patrick C. Shih. 2021. Examining Mobility Among People Living with HIV in Rural Areas. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21, Article 201). Association for Computing Machinery, New York, NY, USA, 1–17.
- [31] René Flores and Edward Telles. 2012. Social stratification in Mexico: Disentangling color, ethnicity, and class. American sociological review 77, 3 (2012), 486–494.
- [32] Marion Fourcade and Kieran Healy. 2017. Categories All the Way Down. Historische Sozialforschung 42 (2017), 286-296.
- [33] Tracy Frey and Roxanne K. Young. 2020. Race/Ethnicity. AMA Manual of Style. https://www.amamanualofstyle.com/view/10.1093/jama/ 9780190246556.001.0001/med-9780190246556-chapter-11-div2-23. Accessed: 2021-8-5.
- [34] Georita M Frierson, David M Williams, Shira Dunsiger, Beth A Lewis, Jessica A Whiteley, Anna E Albrecht, John M Jakicic, Santina M Horowitz, and Bess H Marcus. 2008. Recruitment of a racially and ethnically diverse sample into a physical activity efficacy trial. *Clinical Trials* 5, 5 (2008), 504–516.
- [35] Radhika Garg. 2021. Understanding Tensions and Resilient Practices that Emerge from Technology Use in Asian India Families in the U.S.: The Case of COVID-19. Proc. ACM Hum.-Comput. Interact. 5, CSCW2 (Oct. 2021), 1–33.
- [36] Darren Gergle and Desney S Tan. 2014. Experimental Research in HCI. In Ways of Knowing in HCI, Judith S Olson and Wendy A Kellogg (Eds.). Springer New York, New York, NY, 191–227.
- [37] Jonathan Gordils, Nicolas Sommet, Andrew J Elliot, and Jeremy P Jamieson. 2020. Racial Income Inequality, Perceptions of Competition, and Negative Interracial Outcomes. Social Psychological and Personality Science 11, 1 (Jan. 2020), 74–87.
- [38] Robert M Groves, Eleanor Singer, and Amy Corning. 2000. Leverage-Saliency Theory of Survey Participation: Description and an Illustration. Public Opinion Quarterly 64, 3 (2000), 299–308.
- [39] Maureen T Hallinan. 2001. Sociological Perspectives on Black-White Inequalities in American Schooling. Sociology of Education 74 (2001), 50–70.
- [40] Foad Hamidi, Lydia Stamato, Lisa Scheifele, Rian Ciela Visscher Hammond, and S Nisa Asgarali-Hoffman. 2021. "Turning the Invisible Visible": Transdisciplinary Bioart Explorations in Human-DNA Interaction. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–15.
- [41] Paul H P Hanel and Katia C Vione. 2016. Do Student Samples Provide an Accurate Estimate of the General Public? PLOS One 11, 12 (Dec. 2016), e0168354.
- [42] Alex Hanna, Emily Denton, Andrew Smart, and Jamila Smith-Loud. 2020. Towards a critical race methodology in algorithmic fairness. In Proceedings of the 2020 conference on fairness, accountability, and transparency. 501–512.
- [43] Christina Harrington and Tawanna R Dillahunt. 2021. Eliciting Tech Futures Among Black Young Adults: A Case Study of Remote Speculative Co-Design. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21, Article 397). Association for Computing Machinery, New York, NY, USA, 1–15.
- [44] Christina N Harrington, Shamika Klassen, and Yolanda A Rankin. 2022. "All that You Touch, You Change": Expanding the Canon of Speculative Design Towards Black Futuring. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (New Orleans, LA, USA) (CHI '22, Article 450). Association for Computing Machinery, New York, NY, USA, 1–10.
- [45] Joseph Henrich, Steven J Heine, and Ara Norenzayan. 2010. The weirdest people in the world? Behavioral and Brain Sciences 33, 2-3 (June 2010), 61–83; discussion 83–135.
- [46] P J Henry. 2008. Student Sampling as a Theoretical Problem. Psychological Inquiry 19, 2 (2008), 114-126.
- [47] Julia Himmelsbach, Stephanie Schwarz, Cornelia Gerdenitsch, Beatrix Wais-Zechmann, Jan Bobeth, and Manfred Tscheligi. 2019. Do We Care About Diversity in Human Computer Interaction: A Comprehensive Content Analysis on Diversity Dimensions in Research. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) (CHI '19, Paper 490). Association for Computing Machinery, New York, NY, USA, 1–16.
- [48] Ariel T Holland and Latha P Palaniappan. 2012. Problems with the collection and interpretation of Asian-American health data: omission, aggregation, and extrapolation. Annals of Epidemiology 22, 6 (June 2012), 397–405.
- [49] Catherine Hu, Christopher Perdriau, Christopher Mendez, Caroline Gao, Abrar Fallatah, and Margaret Burnett. 2021. Toward a Socioeconomic-Aware HCI: Five Facets. (Aug. 2021). arXiv:2108.13477 [cs.HC]
- [50] Lily Hu and Issa Kohler-Hausmann. 2020. What's sex got to do with machine learning?. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (Barcelona, Spain) (FAT* '20). Association for Computing Machinery, New York, NY, USA, 513.
- [51] Jennifer L Hughes, Abigail A Camden, and Tenzin Yangchen. 2016. Rethinking and updating demographic questions: Guidance to improve descriptions of research samples. Psi Chi Journal of Psychological Research 21, 3 (2016), 138–151.
- [52] Karen R Humes, Nicholas A Jones, Roberto R Ramirez, and Others. 2011. Overview of race and Hispanic origin: 2010. (2011).
- [53] Institute of Medicine (US) Subcommittee on Standardized Collection of Race/Ethnicity Data for Healthcare Quality Improvement. 2014. Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. National Academies Press (US), Washington (DC).

- [54] Northwestern University Institutional Review Board Office. [n.d.]. Protocol Templates and Forms. https://irb.northwestern.edu/resourcesguidance/protocol-templates-forms/index.html. Accessed: 2022-8-31.
- [55] University of Michigan Institutional Review Board Office. [n.d.]. IRB Application Process. https://research-compliance.umich.edu/irb-applicationprocess. Accessed: 2022-8-31.
- [56] Institutional Review Board Office, Carnegie Mellon University. [n.d.]. Guidance & Forms Office of Research Integrity and Compliance Carnegie Mellon University. https://www.cmu.edu/research-compliance/human-subjects-research/guidance-forms.html. Accessed: 2022-8-13.
- [57] Abigail Z Jacobs and Hanna Wallach. 2019. Measurement and Fairness. arXiv (Dec. 2019). arXiv:1912.05511 [cs.CY]
- [58] Sasha Shen Johfre and Jeremy Freese. 2021. Reconsidering the Reference Category. Sociological Methodology 51, 2 (Aug. 2021), 253–269.
- [59] C P Jones. 2001. Invited commentary: "race," racism, and the practice of epidemiology. American Journal of Epidemiology 154, 4 (Aug. 2001), 299–304; discussion 305–6.
- [60] Camara Phyllis Jones, 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 (2008), 496–504.
- [61] Phoebe K. Chua and Melissa Mazmanian. 2021. What Are You Doing With Your Phone? How Social Class Frames Parent-Teen Tensions around Teens' Smartphone Use. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–12.
- [62] Bliss Kaneshiro, Olga Geling, Kapuaola Gellert, and Lynnae Millar. 2011. The challenges of collecting data on race and ethnicity in a diverse, multiethnic state. *Hawaii Medical Journal* 70, 8 (Aug. 2011), 168–171.
- [63] J S Kaufman. 1999. How inconsistencies in racial classification demystify the race construct in public health statistics. *Epidemiology* 10, 2 (March 1999), 101–103.
- [64] J S Kaufman and R S Cooper. 2001. Commentary: considerations for use of racial/ethnic classification in etiologic research. American Journal of Epidemiology 154, 4 (Aug. 2001), 291–298.
- [65] Lindsey Konkel. 2015. Racial and Ethnic Disparities in Research Studies: The Challenge of Creating More Diverse Cohorts. Environmental Health Perspectives 123, 12 (Dec. 2015), A297–302.
- [66] Masoud Mehrabi Koushki, Borke Obada-Obieh, Jun Ho Huh, and Konstantin Beznosov. 2021. On Smartphone Users' Difficulty with Understanding Implicit Authentication. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–14.
- [67] Nancy Krieger, Jaquelyn L Jahn, and Pamela D Waterman. 2017. Jim Crow and estrogen-receptor-negative breast cancer: US-born black and white non-Hispanic women, 1992-2012. Cancer Causes & Control 28, 1 (Jan. 2017), 49–59.
- [68] Min Kyung Lee and Katherine Rich. 2021. Who Is Included in Human Perceptions of AI?: Trust and Perceived Fairness around Healthcare AI and Cultural Mistrust. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21, Article 138). Association for Computing Machinery, New York, NY, USA, 1–14.
- [69] Sharon M Lee. 1993. Racial classifications in the US census: 1890-1990. Ethnic and racial studies 16, 1 (Jan. 1993), 75-94.
- [70] Kevin E Levay, Jeremy Freese, and James N Druckman. 2016. The Demographic and Political Composition of Mechanical Turk Samples. SAGE Open 6, 1 (Jan. 2016), 2158244016636433.
- [71] ACM Digital Library. [n.d.].
- [72] Sebastian Linxen, Vincent Cassau, and Christian Sturm. 2021. Culture and HCI: A still slowly growing field of research. Findings from a systematic, comparative mapping review. In Proceedings of the XXI International Conference on Human Computer Interaction (Málaga, Spain) (Interacción '21, Article 25). Association for Computing Machinery, New York, NY, USA, 1–5.
- [73] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How WEIRD is CHI?. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21, Article 143). Association for Computing Machinery, New York, NY, USA, 1–14.
- [74] Sarah Lopez, Yi Yang, Kevin Beltran, Soo Jung Kim, Jennifer Cruz Hernandez, Chelsy Simran, Bingkun Yang, and Beste F Yuksel. 2019. Investigating Implicit Gender Bias and Embodiment of White Males in Virtual Reality with Full Body Visuomotor Synchrony. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–12.
- [75] Neda Maghbouleh, Ariela Schachter, and René D Flores. 2022. Middle Eastern and North African Americans may not be perceived, nor perceive themselves, to be White. PNAS 119, 7 (Feb. 2022).
- [76] Patrick L Mason and Andrew Matella. 2014. Stigmatization and racial selection after September 11, 2001: self-identity among Arab and Islamic Americans. IZA Journal of Migration 3, 1 (Oct. 2014), 1–21.
- [77] Omar Mubin, Fady Alnajjar, and Mudassar Arsalan. 2022. HCI Research in the Middle East and North Africa: A Bibliometric and Socioeconomic Overview. International Journal of Human–Computer Interaction 38, 16 (Oct. 2022), 1546–1562.
- [78] Vivek H Murthy, Harlan M Krumholz, and Cary P Gross. 2004. Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA 291, 22 (June 2004), 2720–2726.
- [79] National Institutes of Health. 2001. NOT-OD-01-053: NIH POLICY ON REPORTING RACE AND ETHNICITY DATA: SUBJECTS IN CLINICAL RESEARCH. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-01-053.html. Accessed: 2021-8-5.
- [80] National Science Foundation. 2017. Technical Notes. https://www.nsf.gov/statistics/2017/nsf17310/technical-notes.cfm. Accessed: 2021-8-5.

- [81] Mmachi God'sglory Obiorah, James K L Hammerman, Becky Rother, Will Granger, Haley Margaret West, Michael Horn, and Laura Trouille. 2021. U!Scientist: Designing for People-Powered Research in Museums. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–14.
- [82] Anna Offenwanger, Alan John Milligan, Minsuk Chang, Julia Bullard, and Dongwook Yoon. 2021. Diagnosing Bias in the Gender Representation of HCI Research Participants: How it Happens and Where We Are. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21, Article 399). Association for Computing Machinery, New York, NY, USA, 1–18.
- [83] Ihudiya Finda Ogbonnaya-Ogburu, Angela D R Smith, Alexandra To, and Kentaro Toyama. 2020. Critical Race Theory for HCI. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–16.
- [84] Ihudiya Finda Ogbonnaya-Ogburu, Kentaro Toyama, and Tawanna R Dillahunt. 2019. Towards an Effective Digital Literacy Intervention to Assist Returning Citizens with Job Search. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) (CHI '19, Paper 85). Association for Computing Machinery, New York, NY, USA, 1–12.
- [85] Sam S Oh, Joshua Galanter, Neeta Thakur, Maria Pino-Yanes, Nicolas E Barcelo, Marquitta J White, Danielle M de Bruin, Ruth M Greenblatt, Kirsten Bibbins-Domingo, Alan H B Wu, Luisa N Borrell, Chris Gunter, Neil R Powe, and Esteban G Burchard. 2015. Diversity in Clinical and Biomedical Research: A Promise Yet to Be Fulfilled. *PLOS Medicine* 12, 12 (Dec. 2015), e1001918.
- [86] Newton G Osborne and Marvin D Feit. 1992. The Use of Race in Medical Research. JAMA 267, 2 (Jan. 1992), 275-279.
- [87] Cale J Passmore, Max V Birk, and Regan L Mandryk. 2018. The Privilege of Immersion: Racial and Ethnic Experiences, Perceptions, and Beliefs in Digital Gaming. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18, Paper 383). Association for Computing Machinery, New York, NY, USA, 1–19.
- [88] Pew Research Center. 2015. Multiracial in America: Proud, Diverse and Growing in Numbers. https://www.pewsocialtrends.org/wp-content/ uploads/sites/3/2015/06/2015-06-11_multiracial-in-america_final-updated.pdf. Accessed: 2021-6-25.
- [89] Laura R Pina, Carmen Gonzalez, Carolina Nieto, Wendy Roldan, Edgar Onofre, and Jason C Yip. 2018. How Latino Children in the U.S. Engage in Collaborative Online Information Problem Solving with their Families. Proc. ACM Hum.-Comput. Interact. 2, CSCW (Nov. 2018), 1–26.
- [90] Thomas V Pollet and Tamsin K Saxton. 2019. How Diverse Are the Samples Used in the Journals 'Evolution & Human Behavior' and 'Evolutionary Psychology'? Evolutionary Psychological Science 5, 3 (Sept. 2019), 357–368.
- [91] Race in HCI Collective, Angela D R Smith, Adriana Alvarado Garcia, Ian Arawjo, Audrey Bennett, Khalia Braswell, Bryan Dosono, Ron Eglash, Denae Ford, Daniel Gardner, Shamika Goddard, Jaye Nias, Cale Passmore, Yolanda Rankin, Naba Rizvi, Carol F Scott, Jakita Thomas, Alexandra To, Ihudiya Finda Ogbonnaya-Ogburu, and Marisol Wong-Villacres. 2021. Keepin' it real about race in HCI. Interactions 28, 5 (Aug. 2021), 28–33.
- [92] Shan M Randhawa, Tallal Ahmad, Jay Chen, and Agha Ali Raza. 2021. Karamad: A Voice-based Crowdsourcing Platform for Underserved Populations. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–15.
- [93] Yolanda A Rankin and Na-Eun Han. 2019. Exploring the Plurality of Black Women's Gameplay Experiences. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) (CHI '19, Paper 139). Association for Computing Machinery, New York, NY, USA, 1–12.
- [94] Yolanda A Rankin and Kallayah K Henderson. 2021. Resisting Racism in Tech Design: Centering the Experiences of Black Youth. Proc. ACM Hum.-Comput. Interact. 5, CSCW1 (April 2021), 1–32.
- [95] Yolanda A Rankin and Jakita O Thomas. 2019. Straighten up and fly right: rethinking intersectionality in HCI research. Interactions 26, 6 (Oct. 2019), 64–68.
- [96] REDACTED. 2022. Collecting and Reporting Race and Ethnicity Data in HCI. In Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (New Orleans, LA, USA) (CHI EA '22, Article 327). Association for Computing Machinery, New York, NY, USA, 1–8.
- [97] University of Toronto Research Ethics Boards. [n.d.]. Research Ethics Boards. https://research.utoronto.ca/ethics-human-research/research-ethicsboards. Accessed: 2022-8-31.
- [98] Zarine L Rocha and Peter J Aspinall. 2020. Introduction: Measuring Mixedness Around the World. In The Palgrave International Handbook of Mixed Racial and Ethnic Classification. Springer International Publishing, 1–25.
- [99] Wendy D Roth. 2016. The multiple dimensions of race. Ethnic and Racial Studies 39, 8 (2016), 1310–1338.
- [100] Georgia Robins Sadler, Hau-Chen Lee, Rod Seung-Hwan Lim, and Judith Fullerton. 2010. Recruitment of hard-to-reach population subgroups via adaptations of the snowball sampling strategy. Nursing & Health Sciences 12, 3 (Sept. 2010), 369–374.
- [101] Helen Hatab Samhan. 2001. Who Are Arab Americans?
- [102] Morgan Klaus Scheuerman, Katta Spiel, Oliver L. Haimson, Foad Hamidi, and Stacy M. Branham. 2021. HCI Gender Guidelines. https://www.morganklaus.com/gender-guidelines.html. Accessed: 2021-8-6.
- [103] Dean Schillinger and Urmimala Sarkar. 2009. Numbers don't lie, but do they tell the whole story? Diabetes Care 32, 9 (Sept. 2009), 1746-1747.
- [104] Ari Schlesinger, W Keith Edwards, and Rebecca E Grinter. 2017. Intersectional HCI: Engaging Identity through Gender, Race, and Class. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 5412–5427.
- [105] Jonathan Schwabish and Alice Feng. 2020. Applying Racial Equity Awareness in Data Visualization. (Aug. 2020).
- [106] Patrick Simon, Victor Piche, and Amelie A Gagnon (Eds.). 2015. Social statistics and ethnic diversity: Cross-national perspectives in classifications and identity politics (1 ed.). Springer International Publishing, Cham, Switzerland.

- [107] Angela D R Smith, Alex A Ahmed, Adriana Alvarado Garcia, Bryan Dosono, Ihudiya Ogbonnaya-Ogburu, Yolanda Rankin, Alexandra To, and Kentaro Toyama. 2020. What's Race Got To Do With It? Engaging in Race in HCI. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–8.
- [108] Katta Spiel, Oliver L Haimson, and Danielle Lottridge. 2019. How to do better with gender on surveys: a guide for HCI researchers. Interactions 26, 4 (June 2019), 62–65.
- [109] Simone Stumpf, Anicia Peters, Shaowen Bardzell, Margaret Burnett, Daniela Busse, Jessica Cauchard, and Elizabeth Churchill. 2020. Gender-Inclusive HCI Research and Design: A Conceptual Review. Now Foundations and Trends.
- [110] Christian Sturm, Alice Oh, Sebastian Linxen, Jose Abdelnour Nocera, Susan Dray, and Katharina Reinecke. 2015. How WEIRD is HCI? Extending HCI Principles to other Countries and Cultures. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (Seoul, Republic of Korea) (CHI EA '15). Association for Computing Machinery, New York, NY, USA, 2425–2428.
- [111] Edward Telles. 2014. Pigmentocracies: Ethnicity, race, and color in Latin America. UNC Press Books.
- [112] Kentaro Toyama. 2010. Human-Computer Interaction and Global Development. Foundations and Trends[®] in Human-Computer Interaction 4, 1 (2010), 1–79.
- [113] US Census Bureau. [n.d.]. 2020 Census Frequently Asked Questions About Race and Ethnicity. https://www.census.gov/programs-surveys/decennialcensus/decade/2020/planning-management/release/faqs-race-ethnicity.html. Accessed: 2022-8-5.
- [114] US Census Bureau. 2019. National Demographic Analysis Tables: 2020. Accessed: 2021-3-8.
- [115] Tavish Vaidya, Daniel Votipka, Michelle L Mazurek, and Micah Sherr. 2019. Does Being Verified Make You More Credible? Account Verification's Effect on Tweet Credibility. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–13.
- [116] Judy van Biljon and Karen Renaud. 2019. Human-Computer Interaction for Development (HCI4D): The Southern African Landscape. In Information and Communication Technologies for Development. Strengthening Southern-Driven Cooperation as a Catalyst for ICT4D. Springer International Publishing, 253–266.
- [117] Kelly Walters, Dimitri A Christakis, and Davene R Wright. 2018. Are Mechanical Turk worker samples representative of health status and health behaviors in the U.S.? PLOS One 13, 6 (June 2018), e0198835.
- [118] Connie Wanberg, Gokce Basbug, Edwin A J Van Hooft, and Archana Samtani. 2012. Navigating the black hole: Explicating layers of job search context and adaptational responses. *Personnel Psychology* 65, 4 (Dec. 2012), 887–926.
- [119] Michael Weiner. 2022. Routledge Handbook of Race and Ethnicity in Asia. Routledge.
- [120] Earnest Wheeler and Tawanna R Dillahunt. 2018. Navigating the Job Search as a Low-Resourced Job Seeker. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18, Paper 48). Association for Computing Machinery, New York, NY, USA, 1–10.
- [121] Kim M Williams. 2006. Mark one or more. University of Michigan Press.
- [122] Marisol Wong-Villacres, Adriana Alvarado Garcia, Karla Badillo-Urquiola, Mayra Donaji Barrera Machuca, Marianela Ciolfi Felice, Laura S Gaytán-Lugo, Oscar A Lemus, Pedro Reynolds-Cuéllar, and Monica Perusquía-Hernández. 2021. Lessons from Latin America: embracing horizontality to reconstruct HCI as a pluriverse. *Interactions* 28, 2 (March 2021), 56–63.
- [123] Sarita Yardi and Amy Bruckman. 2012. Income, race, and class: exploring socioeconomic differences in family technology use. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Austin, Texas, USA) (CHI '12). Association for Computing Machinery, New York, NY, USA, 3041–3050.
- [124] Henri-Michel Yéré, Mavis Machirori, and Jantina De Vries. 2022. Unpacking race and ethnicity in African genomics research. Nature reviews. Genetics 23, 8 (Aug. 2022), 455–456.