A Decade of Demographics in Computing Education Research: A Critical Review of Trends in Collection, Reporting, and Use

Alannah Oleson*
Benjamin Xie*
olesona@uw.edu
bxie@uw.edu
University of Washington
Seattle, Washington, USA

Jean Salac
University of Washington
Seattle, Washington, USA
salac@uw.edu

Jayne Everson
University of Washington
Seattle, Washington, USA
everjay@uw.edu

F. Megumi Kivuva
Bard College
Annandale-on-Hudson, New York
USA
mk5730@bard.edu

Amy J. Ko
University of Washington
Seattle, Washington, USA
ajko@uw.edu

ABSTRACT

Computing education research (CER) has used demographic data to understand learners' identities, backgrounds, and contexts for efforts such as culturally-responsive computing. Prior work indicates that failing to elucidate and critically engage with the implicit assumptions of a field can unintentionally reinforce power structures that further marginalize people from non-dominant groups. The goal of this paper is two-fold: to understand what populations CER researchers have studied, and to surface implicit assumptions about how researchers have collected, reported, and used demographic data on these populations. We conducted a content analysis of 510 peer-reviewed papers published in 12 CER venues from 2012 to 2021. We found that (1) 60% of papers studied older learners in formal contexts (i.e. post-secondary education); (2) 68% of papers left unclear how researchers collected demographic data; and (3) while 94% of papers were single-site studies, only 14% addressed the limitations of their contexts. We also identified hegemonic norms through ambiguous aggregate term usage (e.g. underrepresented, diverse) in 23% of papers, and through incomplete reporting of demographics (i.e. leaving out demographics for some participants in their sample) in 35% of papers. We discuss the implications of these findings for the CER field, raising considerations for CER researchers to keep in mind when collecting, reporting, and using demographic data.

CCS CONCEPTS

• **Social and professional topics → Computing education**.

*Both authors contributed equally to this research.

ICER 2022, August 7–11, 2022, Lugano and Virtual Event, Switzerland

ACM ISBN 978-1-4503-9194-8/22/08.
https://doi.org/10.1145/3501385.3543967

KEYWORDS

demographic data, content analysis, critical demography, literature review

ACM Reference Format:
Alannah Oleson, Benjamin Xie, Jean Salac, Jayne Everson, F. Megumi Kivuva, and Amy J. Ko. 2022. A Decade of Demographics in Computing Education Research: A Critical Review of Trends in Collection, Reporting, and Use. In *Proceedings of the 2022 ACM Conference on International Computing Education Research V.1 (ICER 2022), August 7–11, 2022, Lugano and Virtual Event, Switzerland*. ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3501385.3543967

1 INTRODUCTION

Computing education research (CER) has used demographic data[1] to understand learners' identities, backgrounds, and contexts for numerous diversity, equity, and inclusion efforts. Demographic data can identify disparities that hinder participation [175], such as differences in access, retention, and achievement by gender and ethnicity [62, 210]. Demographics can also illuminate how instructional design differentially impacts populations, in opportunities, preparatory privilege, and prevailing attitudes [78, 79, 143]. Recent efforts have used demographics to consider *intersectional* identities [48] for culturally-responsive learning [141, 153], such as work training for Black men and women [114], transformative justice programs for Black and Latina girls [61], and electronic textiles with American-Indian boys [198].

CER researchers' *choices* of how to gather and use demographics shape our understandings of learners and teachers, impacting its future reporting and use [67]. For example, the decision to collect gender as a binary construct (e.g. woman, man) has resulted in systemic erasure of non-binary learners [158, 230]. Reporting demographics in CER publications involves considering how the data was collected, perceptions of what audiences will deem valuable, well-being of participants, and pragmatic constraints like page lengths [10, 67]. Finally, how we use demographics in studies affects how others interpret and build off of findings. Researchers have used

[1]"Demographics" and "demographic data" are distinct concepts, but we use "demographics" for brevity.

gender data, for example, to describe participants as part of analysis (e.g. [30, 183, 223]), and to identify limitations (e.g. [149, 199]). All of these decisions impact CER's collective understanding of how participants' contexts can impact access, engagement, and achievement.

Demographic decisions are not made in isolation: They are influenced by broader and intersecting community norms [102, 233]. Cultural and contextual norms can also impact decisions about demographics. A United States (US)-based research team might decide to report age with US-centric terminology (e.g. "high school"), making interpretation more difficult for readers elsewhere. In contrast, a multi-national research team may decide to explicitly describe differences in grades of school placement across countries relative to age [65], better supporting interpretation. Because most CER contributions undergo scrutiny from others in peer-review processes, implicit community norms also impact decisions around demographics. Prior work suggests CER papers often do not describe demographics in sufficient detail for replication [93], perhaps due to page limits [1]. The norms of government, non-profit, and industry funding sources can also influence decisions around demographics. For example, given women's disproportionately low participation in computing [21, 144], it is unsurprising that funding agencies (e.g. UK Research and Innovation [3], US National Science Foundation [2]) would promote programs that foster women's participation in computing. However, implicit in this definition is the assumption that women are the *only* gender marginalized in computing. While well-meaning, such efforts unintentionally uphold hegemonic norms of gender.

Partly because CER is global [137], the norms that shape decisions around demographics in CER are often ambiguous, inconsistent, and not fully understood. This makes them difficult to directly critique for purposes of fostering more rigorous research that future work can build upon [93] and more critical research that fosters more just computing communities [121]. While research communities and prior work have defined recommendations on reporting demographic [9, 151], it is unclear how closely researchers follow these recommendations.

In this paper, we attempt to identify CER's emerging demographics norms, asking:

(1) What *populations* have CER papers studied?
(2) How have demographics been *collected* in CER papers?
(3) What kinds of demographics have been *reported* in CER papers and what kind of language do authors use when reporting?
(4) How have demographics been *used* in CER papers?

To answer these questions, we applied content analysis [163] on a stratified random sample of 510 peer-reviewed papers published in 2012-2021 in 12 CER journals, conferences, and working groups. Our work builds on prior work by 1) conducting a more *comprehensive* analysis of CER papers to identify demographics norms, and 2) critiquing these norms relative to CER goals of rigor and criticality, and 3) considering the entire "pipeline" of data collection, reporting, and use. From this, we inferred norms of reporting demographic attributes. We then critique these norms to consider what norms on demographics *should be* relative to goals of conducting rigorous and critical research.

2 BACKGROUND

Categorization is a form of abstraction that allows people to interpret large amounts of data through reduction to each item's most salient or relevant characteristics. Bowker and Star define three properties of a classification system: consistent, unique principles for sorting; mutually exclusive categories; and complete coverage of what items are or can be [37]. However, classification is a value-laden activity. Because it is a reduction of richness for abstraction, the process implies some information loss. The choices of what information to capture and leave out, as well as how to represent captured data, are *design decisions*, and like all design decisions, these choices embed the values and biases of those who make them (intentionally or not) [70].

Demographics are, at a high level, labels for categories of people, reducing identity for quantification and analysis [233] through a process of assigning people to groups that distinguish them from each other [67]. This makes it difficult to find a demographic classification that works for all contexts, purposes, and peoples. Static, literal, and rigidly-bounded demographic schemes function well only when a user's identify fits into the allowable bounds of the system. *Dominant* groups, those that are privileged [226], unstigmatized [187], and generally favored by social, economic, political, and educational institutions [59, 142] typically design these schemes. These schemes therefore tend to only serve people from dominant populations well, embedding power imbalances and hegemonic norms of the context [99]. For instance, the American Anthropological Association tried (and failed) to eliminate usage of the term "race" from the US Census, asserting that the concept of race was scientifically unsound (as it was developed for discriminatory reasons), and that ethnicity was more accurate descriptor for classification of groups of people [8]. In contrast to dominant groups are *marginalized* groups, those who are not positively privileged or favored and often stigmatized. Prior work shows that most demographic classification schemes created by dominant groups erase the presence of marginalized identities [27, 45], especially racial and ethnic identities, for which there is no apolitical classification scheme [213, 233]. Conceiving of identity as intersectional [17, 48, 182, 197] breaks many demographic classification systems, in that identities can no longer be fully represented by a single (or even a set of) mutually exclusive categories. Furthermore, identity is not static. It is often difficult for classification schemes to account for marginalized [86, 119] and changing [85] identities in a way that authentically represents and respects them.

The field of demography investigates the use of demographics, statistically characterizing populations in different ways. Similar to many quantitative, positivist fields, *conventional* demography assumes the objectivity and independence of demographics, the processes that produce them, and the people involved, thereby ignoring or implicitly accepting norms that reflect a status-quo [173, 191]. This can result in ignoring or misunderstanding the broader consequences of social phenomena, such as how civil rights movements influenced demographic collection and reporting methods [102, 107, 167]. In contrast, *critical* demography enables reflection on the state of demographics and the process that produces them [102, 233]. This paradigm enables the articulation of social, economic, and political context within which demography occurs. This examination of power relationships within statistical data provides a more holistic understanding of not only how populations are categorized, but *why* those particular classifications are used and *how* the given groupings reinforce or challenge existing norms. Critical demography requires consideration of how researcher positionality, and political and theoretical ideas affect interpretation of discoveries [233].

Within CER, recommendations on how to report demographics exist, but they tend to be too high level or incomplete to help us understand norms of demographic data. For example, the American Educational Research Association standards on reporting empirical research mentions the reporting of demographics, but detail on what to report is lacking [9]. Prior CER literature reviews suggest that demographics for students should include ages, education levels, gender, race/ethnicity, prior experience, and regional location [53, 93, 151], but it is unclear how closely CER papers follow these recommendations.

Shortcomings in reporting demographics can hinder the *rigor* of empirical findings in CER, or how papers enable future work to build off them for replication, meta-analysis, and theory building [9, 93]. Heckman et al. conducted a systematic literature review to understand norms of reporting empirical studies, finding that most CER papers only weakly supported replication because they lacked details about participants [93]. Margulieux et al. found similar, with only 49% of the 197 reviewed papers reporting the "basic" demographics of gender (35% of sample, 69 papers), age (21%, 41 papers), prior experience in computing (18%, 35 papers), and race (14%, 28 papers) [145]. A review of pre-college computing activities by McGill et al. found that many of the 92 reviewed papers failed to report important demographics, including socioeconomic status (13%) [151]. Another review of 76 studies applying educational data mining and/or learning analytics techniques for computing education identified that most studies did not collect or report demographic information, potentially leading to confounds [108]. Collectively, these meta-analyses identified how the CER community fails to rigorously report demographics [93].

Recent work in CER has called for more *critical* investigation of demographics that consider existing and historical power structures. Convertino identified how oversimplifying the narrative that women were an underrepresented, invisible monolith is an unproductive reduction, like how women of color in CS resist the dominant discourse of underrepresentation [44, 212]. Ross et al. conducted a more intersectional analysis of survey data, comparing experiences of computing students who were Black women, non-Black women, and Black men to surface the intersection of being Black and being women [189]. Lunn et al. analyzed intersectional demographics with historical context analysis to describe the political, economic, and social factors that may have impacted experiences of women, Black, Hispanic/Latinx and Native American groups in computing [138]. While these prior studies focused on intersectionality across race/ethnicity and gender, Pournaghshband & Medel called for intersectional approaches that went beyond these two dimensions [174]. Elements of social identity they highlighted included gender, race, socioeconomic status (SES), geographic location, ablebodiedness, culture, sexual orientation, and linguistic background. Collectively, these papers highlight the need to consider multiple dimensions of demographics that go beyond considering gender or race/ethnicity in isolation.

3 METHOD

We use the lens of *critical* demography to explore how CER research reports 11 demographics, including the use of aggregate terms. By emphasizing how implicit CER norms guide our data collection, reporting, and usage, we sought to contribute to broader conversations about justice, equity, and power around the teaching and learning of computing (c.f. [121, 170]), seeking to identify and abolish hegemonic norms that contribute to further marginalization [83, 84, 163].

We surfaced norms through a content analysis of 510 peer-reviewed papers from 12 CER venues. Content analysis summarizes content (e.g. written text) systematically [163]. It involves selecting content to analyze, defining units of analysis, developing rules for qualitative coding, coding the content, and analyzing the results [139, 163]. This enables description and inferences about the creators, context, audience of the content. By systematically analyzing durable data through a customizable process, content analysis affords transparency, replicability, and flexibility [163, 207].

Prior CER work has used content analysis to examine the thematic landscape of the field [169], student difficulties [155, 171], and a pedagogical content knowledge model [219]. Prior work has also applied content analysis through a critical lens, such as exploring Black women's experience in computing [212] and identifying power structures that reinforce social differences along class, gender, and race [83, 84].

3.1 Dataset: Publications in 12 CER venues from 2012-21

Table 1 outlines the number of papers in each venue at each analysis step. We first downloaded references for 3,429 papers for our *corpus* of papers. We then extracted a random *stratified sample* of 705 papers to consider. Finally, we checked the stratified sample against our inclusion criteria, leaving 510 papers for content analysis.

The 12 CER venues reflected those included in prior literature reviews (e.g. ICER, ITiCSE, SIGCSE, TOCE, CSE, Koli) [93, 108, 137, 145, 151, 158], smaller venues (WiPSCE, CSERC), working groups (CompEd WG, ITiCSE WG), and newer venues (CompEd, RESPECT). Table 1 shows the number of years each venue published from 2012-2021 and summary statistics on the number of papers published per year. These summary statistics do *not* consider years where venues published no papers.

We downloaded 3,429 paper references from these venues in January 2022. We used the ACM Digital Library to download papers for the ten ACM-affiliated venues, filtering by content type ("research articles"[2]) and publication date (2012-2021). We filtered using the same dates for CSE papers in Taylor & Francis Online and RESPECT papers in IEEE Xplore, but could not filter by content type, resulting in lower inclusion rates for RESPECT and CSE.

[2] "Research articles" in the ACM Digital Library include both standard research papers and "experience report" formats from venues such as SIGCSE. Our analysis treated both these formats equally, since there are no commonly agreed-upon standards for what constitutes an experience report versus a standard research article.

Table 1: Number of papers downloaded, sampled, and included in our content analysis by venue. **: counts for these venues are not comparable to other venues because content was not downloaded from the ACM Digital Library.

Venue	Num years w/ papers, 2012-21 (max 10)	Median papers/ yr [range]	Corpus (%: venue / total corpus)	Stratified sample (%: venue / total strat. sample)	Included papers (%: venue / total incl. papers)	% that met inclusion criteria (%: incl. papers / strat. sample)
CompEd	1	33	33 (1%)	33 (5%)	30 (6%)	91%
CompEd WG	1	1	1 (0%)	1 (0%)	1 (0%)	100%
CSE**	10	18 [11-39]	201 (6%)	39 (6%)	24 (5%)	62%
CSERC	7	8 [5-18]	68 (2%)	19 (3%)	14 (3%)	74%
ICER	10	26.5 [15-30]	251 (7%)	57 (8%)	46 (9%)	81%
ITiCSE	10	58 [49-84]	612 (18%)	117 (17%)	97 (19%)	83%
ITICSE WG	7	7 [3-9]	43 (1%)	17 (2%)	5 (1%)	29%
Koli	10	20 [12-29]	196 (6%)	43 (6%)	35 (7%)	81%
RESPECT**	6	47.5 [30-85]	313 (9%)	96 (14%)	33 (6%)	34%
SIGCSE	10	110 [105-171]	1,306 (38%)	208 (30%)	169 (33%)	81%
TOCE	10	22.5 [16-49]	257 (7%)	47 (7%)	37 (7%)	79%
WiPSCE	10	11 [8-28]	148 (4%)	28 (4%)	19 (4%)	68%
Total		362.5 [270-446]	3,429 (100%)	705 (100%)	510 (100%)	72%

We then loaded references into RStudio(v3.6.2), extracted publication year and venue, and then randomized order of the rows[3]. We then created unique keys of the form *[publication year]-[venue]-[number]*, with a unique number within a publication year and venue.

We randomly sampled 705 papers, stratified by venue and year. Sampling is common in content analysis [163] and has been used in prior CER content analyses [179, 180, 194]. In 2012-21, some venues aimed to diversify perspectives in CER by focusing on equity and justice (RESPECT) or new regions of the world (CompEd). COVID-19 also canceled some conferences. To ensure these venues were still well-represented in our dataset, we oversampled for them by considering the median number of publications per year to stratify by venue *only for years where there was at least one publication* (Table 1). With a goal sample size of 500 (≈15% of corpus) and estimating that 70% of papers would be pass our inclusion criteria, we used these medians to produce a random sample stratified by venue and year.

3.2 Inclusion & exclusion criteria: Peer-reviewed papers with human participants

We analyzed only peer-reviewed papers to better reflect community norms. Peer-review requires 2-3 community reviewers who are not conflicted [5] to engage with a paper, providing multiple perspectives on what constitutes "acceptable" CER work. This excluded content like panels and posters. We included empirical studies that described human participants because they are the primary source of demographics. operationalizing this to determining that the population of study is human. This criteria was intentionally broad to account for papers with human evaluation but without demographics for their sample. Empirical studies of human-created artifacts (e.g. code snapshots) were included if participants were described with one of the demographics we coded for. This criteria excluded meta-analyses and literature reviews. Our final inclusion criteria, that papers must be written in English (the only language that the entire research team was fluent in) was met due to venue conventions.

3.3 Analysis: Inductive Coding & Thematic Analysis

We analyzed content about demographics in the 510 papers from our stratified sample that met our inclusion criteria.

3.3.1 Collaboratively developing inductive codesets. We developed our codebook through inductive coding, letting the data guide our analysis of themes [208]. We identified themes by analyzing the most-cited paper for each year that fit our inclusion criteria for eight venues: SIGCSE, ICER, ITiCSE, CompEd, TOCE, RESPECT, CSE, and Koli Calling. Each team member analyzed a subset of these papers and noted all demographics these papers reported, and how that data was collected and used. The team then discussed their initial findings, noting high-level trends and emergent themes which formed our initial codesets.

After consulting with a critical data scholar to refine our codesets, we met for two more practice coding rounds. First, we randomly selected three papers from our dataset and coded them simultaneously and then discussed whether or not to apply each code, coming to consensus, and adjusting the code definitions as needed. Then, we asynchronously coded five more random papers each, then met to discuss. By the end of the two practice rounds, team members who would participate in coding felt confident in reliably coding papers.

Codes for populations of study (RQ1, Table 2), collection methods (RQ2, Table 3), and usage patterns (RQ4, Table 5) were each coded dichotomously (present/not present), allowing for multiple codes

[3]Used `dplyr::sample_n()`(v1.0.4). Seed:15.

Table 2: Population of study codeset for classifying study participants (RQ1). Papers with no human participants (no codes in this set) were excluded from further analysis.

Code		Description
Young learners	Formal	Pre-K, primary, and secondary aged students in formal learning contexts (e.g. schools)
	Informal	Pre-K, primary, and secondary aged students in informal learning contexts (e.g. workshops)
	Other	Young learners not covered by the above categories
Older learners	Formal	Post-secondary aged students in formal learning contexts (e.g. universities)
	Informal	Post-secondary aged students in informal learning contexts (e.g. MOOCs)
	Professional	Post-secondary aged students in professional training contexts (e.g. coding bootcamps)
	Other	Older learners not covered by the above categories
Educators	Formal primary, secondary	Educators in Pre-K, primary, and secondary formal learning contexts (e.g. school teachers)
	Informal primary, secondary	Educators in Pre-K, primary, and secondary informal learning contexts (e.g. workshop leaders)
	Post-secondary	Educators in post-secondary learning contexts (e.g. university lecturers)
	Professional	Educators in professional training contexts (e.g. teacher education)
	Other	Educators not covered by the above categories
Professionals	Computing	Those working in technology-related jobs (e.g. software designers)
	Non-computing	Those working in jobs outside the technology sector (e.g. medical professionals)
Other / Unsure		Population of study that does not fit the above categories, OR some ambiguity prevents full identification of the population

Table 3: Demographics collection method codeset for understanding how CER papers obtained demographic data (RQ2).

Code	Description
Self report: Existing	Asks participants to self-report demographics, using some referenced existing instrument. If marked, we captured the instrument.
Self report: Custom	Asks participants to self-report demographics, using a custom instrument created by the authors for use in the specific study.
Pre-existing data	Participant demographics are drawn from some pre-existing data source (e.g. admission applications). If marked, we captured the data source.
Reported by another	Participant demographics reported by someone other than participants (e.g. parents). If marked, we captured who reported the data.
Other	Participant demographics was collected in a specified way not covered by the above categories.
Unclear / No mention	Given only the information in the paper, it is unclear how (at least some) demographics were collected.

Table 4: Demographics reported codeset for understanding the categories CER papers used to classify participants (RQ3).

Code	Description: Terms such as ...
Gender	Women, men, non-binary, etc.
Race/Ethnicity	Black, Indigenous, Hispanic, etc.
Nationality	American, international, citizens, etc.
Fluency	English language learner (ELL), German, Tamil, etc.
Ability	Blind, deaf, “special education", etc.
Age/Grade	10-14 years, 12th grade, second-year undergraduates, etc.
Socioeconomic status (SES)	Income, financial aid, free or reduced lunch, etc.
Other household demographics	Parent education, computer use, first-generation, etc.
Geographic location	Rural/urban contexts; locations within countries; “University of X",etc.
Major/Program	Computer science, STEM, “non-computing", etc.
Aggregate term used	Uses an aggregate term for a group of people suggesting proportionality or power relations, e.g. under-represented. May or may not be disaggregated (disaggregation is captured through the above codes).

per paper. The types of demographics reported codeset (RQ3, Table 4) also allowed multiple codes. Aggregate term usage was coded dichotomously, but the rest of the terms trichotomously:

- *yes-fully*: paper fully reported a demographic for all participants. For instance, if a study's sample size was 40 teachers, the paper might report teachers' genders as 15 women, 15 men, and 10 non-binary teachers (15+15+10=40).

Table 5: Demographics usage codeset for understanding how CER papers used demographic data within their projects (RQ4).

Code	Description
Motivation	Motivating the study using demographic-related arguments, e.g. studying experiences of a particular demographic group. Includes using aggregate terms like "under-represented" to motivate.
Description	Describing individuals who directly participated in a study, e.g. students in a CS course.
Contextualization	Describing the broader contexts of a study, e.g. describing school-level demographics for a classroom study.
Analysis	Using demographics as a variable during analysis, e.g. comparing gender differences.
Validity	Justifying the representativeness of a sample or acknowledging demographic-related limitations of the study.
Other	Using demographics in a way not covered by the above categories, or use is ambiguous.
N/A	No demographics were reported. Indicates that no codes were marked for RQ3 (types of demographics reported).

- *yes-incomplete*: paper reported a demographic for some but not all participants. For instance, if a study's sample size was 300 students, the paper might report that their sample contained 50 Black students and 50 Hispanic students, but no information about the other 200 students. We added *yes-incomplete* to better understand the ways that incomplete reporting might interact with hegemonic norms (i.e. what the unspoken "default" categories were implied to be).
- *not-at-all*: paper did not report a demographic for any participants.

3.3.2 Coding CER paper content & post-hoc analysis of trends. The first five authors participated in the coding process, each coding 50-207 papers. Coders individually analyzed each assigned paper, leveraging the understanding of the codesets built through collaborative development, refinement, and practice. To authentically represent our dataset (CER publications), we adhered to a rule of "taking the paper literally" in that our unit of analysis was text, figures, and tables within the body of the paper and attached appendices. This meant that for each positive instance of a code, we could identify a specific phrase within the paper that directly supported our interpretation.

We chose to uphold the perspective on qualitative coding from Hammer and Berland [88], treating codes as an organizational aid to identify themes within our dataset. Accordingly, we did not capture agreement metrics (e.g. inter-rater reliability) between coders, preferring instead to utilize a consensus-based model to resolve uncertainties. When a coder was unsure whether a code applied to a particular paper, they reached out to another author. The two authors reviewed the paper, agreed upon a code, and refined coding rules when necessary. Once the initial coding pass was finished, we performed post-hoc thematic analyses [172] on the coding results to surface broader trends.

3.4 Author positionality

In a paper that explores the nuances of representing people through demographics, it is important for our research team to recognize our own positionality [71] and how our backgrounds may have influenced our values and assumptions. We also recognize the tensions described by Liang et al. [133] in that disclosure of certain identities (especially minoritized identities) can carry social consequences within the research community, and agree that no researcher should feel like they have to individually out themselves or their situations to participate in research. As a result, we choose to report the research team's positionality collectively rather than individually.

Below, we describe some self-reported facets of the team's background using our own codeset developed inductively from our content analysis of CER papers (Table 4). We did this both to engage more deeply with our own analyses, but also to illustrate some ways in which traditional demographic collection may not suffice for understanding a person's identities and values, and especially falls short in supporting intersectional understandings. We invite readers to reflect upon the insights they can and cannot glean from this list of demographics and to apply those reflections in their own work.

- **Gender:** man, non-binary, queer, queer trans woman, woman
- **Race / Ethnicity:** Asian-American, Black, Danish and Chinese, Filipina, white
- **Nationality:** Filipina, Kenyan, USA
- **Fluency:** English, multilingual, Swahili
- **Ability:** Chronic pain, minor physical disabilities, neurodiverse, not disabled
- **Age / Grade:** Graduate student, post-PhD, 22, 25, 41
- **Socioeconomic status (SES):** financially stable, low-income, rent-burdened, upper-middle class
- **Other family/household:** grew up low income, immigrant, immigrant mother
- **Geographic location:** New York, Northwest state in the US, Pacific Northwest US, Washington
- **Major / Program of Study:** Computer Science, Experimental Humanities, Information Science, Spanish
- **Aggregate terms:** BIPOC, first-generation, CS major, LGBTQ+, previously rural, privileged, underrepresented

We additionally emphasize that we were situated in U.S.-centric contexts. This influenced our qualitative analysis, terminology, and values throughout this research, likely biasing analysis and reporting in U.S.-centric ways.

4 RESULTS

Our goal is to understand broader trends around demographics in CER, and how we, as a community, can be more mindful of how we collect, report, and use demographics. These trends reflect CER community *norms*. Not all norms are hegemonic, nor do all papers follow these norms. In the spirit of critical generosity [128], we do not directly cite most papers to preserve anonymity. Instead, at the end of each subsection, we directly cite what we consider to be exemplars of demographic collection, reporting, and usage. These exemplars are not all-encompassing.

Subsections roughly adhere to the following structure: We first introduce broad related work from adjacent fields, followed by related work from CER. We next present results from reviewed papers and post-hoc analysis, and close with exemplar papers. Unless otherwise stated, proportions presented are based on the 510 analyzed papers.

4.1 RQ1: Populations of Study

The papers we analyzed studied various populations (Figure 1). Most papers (60%; 304) studied older learners in formal settings, e.g. post-secondary learners in a university course. Formal settings were also the most common for young learners, accounting for 16% (80) of analyzed papers. Notably, while studies on older learners dwarfed studies on young learners in formal settings, the reverse was true for formal *educators*, with 13% (68) investigating primary and secondary educators and only 4% (19) investigating post-secondary educators.

Most papers (86%; 441) only studied one population, but 57 papers (11%) studied two, and 12 (2%) studied three or more. The most common multi-population studies (26%; 18/69 papers) examined both students and teachers in primary or secondary schools (e.g. [20, 109, 186, 217]). Others investigated both learners (e.g. [28, 54]) and educators (e.g. [55, 157]) across formal and informal learning contexts. Some analyzed both young and older learners in informal settings (e.g. [35, 80, 91]).

4.2 RQ2: How Papers *Collected* Demographics

Heckman et al. found that 76% of CER papers in their sample utilized only one type of data source, with surveys as the sole source in 30% of all papers [93]. We build upon their work by characterizing *who* provided the data and *how* instruments were created.

Figure 2 shows the results of our deductive coding of how CER papers collected demographics. Participant self-report was the most common way to collect demographics in analyzed papers. Similar to [93], 29% (147) created custom instruments for their studies. However, researchers rarely described their custom instruments sufficiently for replication. Only 3% (14) of papers used existing instruments to collect demographics. Oftentimes these were surveys created by established organizations, like the US Computer Science Teachers' Association [49] and National Center for Women & Information Technology [161].

Another 5% (23) of papers used pre-existing datasets, mostly (16/23) relying on enrollment data at their institution (e.g. *2019-iticse-0016*). Other preexisting datasets included applications for educational opportunities (e.g. *2017-toce-0004*), the Computing Research Association Taulbee survey [234] and US census data. Using existing datasets affected reporting of some demographics. For example, *2021-icer-0007* acknowledged how their university's registrar data limited their analysis to a binary gender classification (Male/Female). Reliance on pre-existing datasets sometimes required reduction of demographics, typically conforming with existing norms.

In 9% (46) of papers, demographics of participants was reported by another party (Figure 2). In most papers, it was authors reporting geographic locations (e.g. *2019-comped-002*). Educators also reported demographics like students' ages and abilities (*2013-cse-0001*). However, having instructors report demographics resulted in some reductions. For example, *2021-iticse-0011* relied on instructors using names and photos to classify students as male or female. For young learners in formal contexts, teachers and/or parents reported demographics like grade, gender, race or ethnicity, and family information (e.g. *2018-sigcse-0026*).

Most papers (68%; 346) did not provide sufficient information to determine data collection (*unclear* in Figure 2). This trend was problematic because knowing how data was collected is critical in the validity of its reporting and use. For papers using preexisting data or relying on reporting by another stakeholder, collection techniques can introduce reductions (e.g. eligibility for free or reduced lunch as proxy for family's socioeconomic status), non-consensual representation (unclear if participants consented to collection, reporting, and use of demographics), or biases (e.g. teachers using photos and names to determine binary gender introduces stereotype threat [205]).

Exemplary papers demonstrated robust descriptions of researchers collecting demographics in *justified*, *transparent*, and *responsible* ways. Sharmin et al. [200] described data collection across 3 surveys, what surveys collected which demographics, and how they used the Computer Programming Self-Efficacy Scale (CPSES) [178]. Cutts et al. also demonstrated transparent and justified demographics collection, including a table summarizing the source, time and location of data collection, method of collection, number of responses, and the purpose of collection [50]. McGee et al. signaled responsible research practices by explicitly mentioning that they collected demographics through a data sharing agreement with a public school system [150].

4.3 RQ3: How CER Papers *Reported* Demographics

4.3.1 Gender. Gender is a social construction of an identity facet, not an innate biological quality [174]. When gender and sex are conflated, gender is often framed as binary, immutable, and physiological, even though these perspectives are largely unfounded [119]. Nonetheless, these conceptions abound in computing research, often erasing the existence of non-binary and transgender individuals [117].

Gender is a popular demographic to collect in CER, partially due to the explicit focus on women in broadening participation efforts (e.g. [160]). A literature review of computing outreach activities found that 72.5% (58/80) of papers reported gender [53]. Another review of CER papers in 2000-2005 found that gender was the most common mediating/moderating variable [179, 180], despite claims from statisticians that gender and other attributes of identity are not explanatory variables [101, 228]. Other CER-related investigations of gender involve analyses of equity efforts, such as faculty perspectives on BPC efforts targeted at women [87] and a review of strategies to support womens' participation in computing education [158]. Unfortunately, similar to broader discussions of gender belonging, "gender diversity" in CER is too

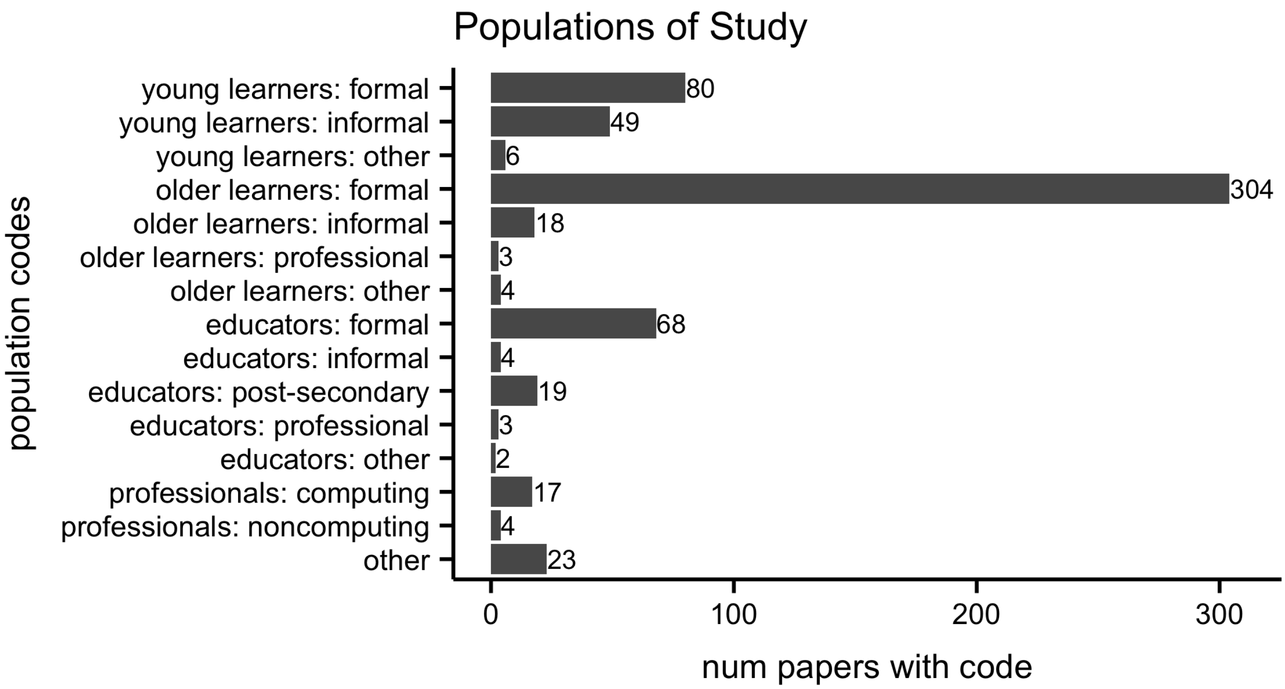

Figure 1: RQ1: Codes reflecting the frequency at which analyzed papers studied different populations. Total number of codes (604) exceeds the number of papers analyzed because 69 papers studied multiple populations.

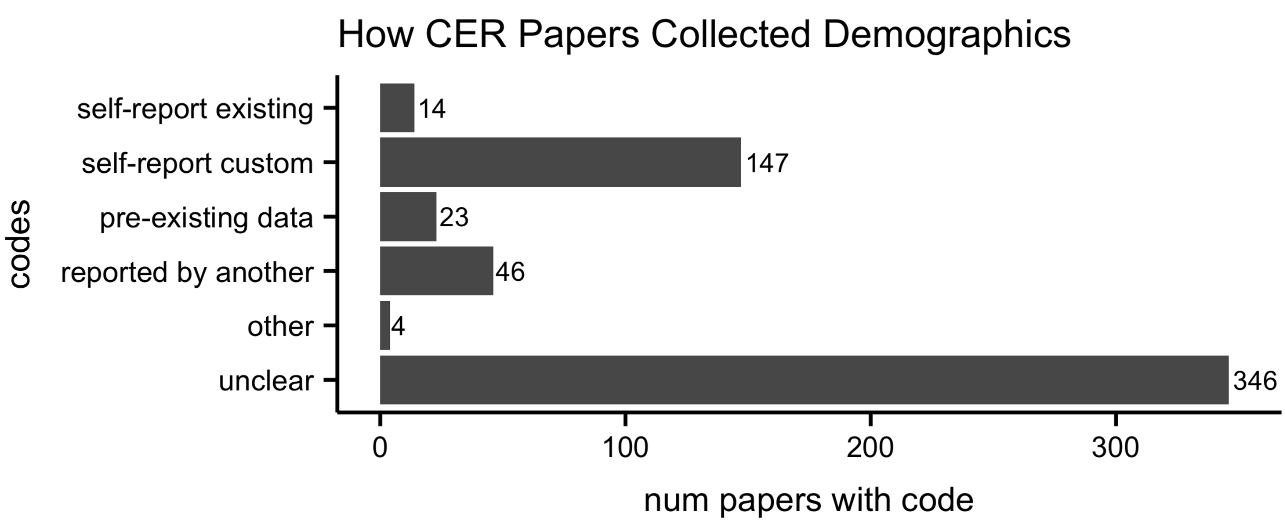

Figure 2: RQ2 Results: How papers collected demographics. Total number of codes (580) exceeds the number of papers analyzed because 65 papers collected demographics 2-3 ways.

often reduced to only the inclusion of women and girls. Erasure of non-binary genders may be due to small sample sizes (c.f. [15]), and even when studies include data from non-binary students, biases in self-reported demographics may limit findings [230]. Erasure may also be systemic: The current version of ACM Computing Classification System (CCS) for describing paper content only includes men and women under the "Gender" subtree (Social and professional topics → User characteristics → Gender) [4], precluding accurate classification of works focusing on other marginalized genders. Pournaghshband and Medal recently called for more intersectional conceptions of identity in CS pedagogies, arguing for a "non-binary aware" approach to demographic collection to avoid

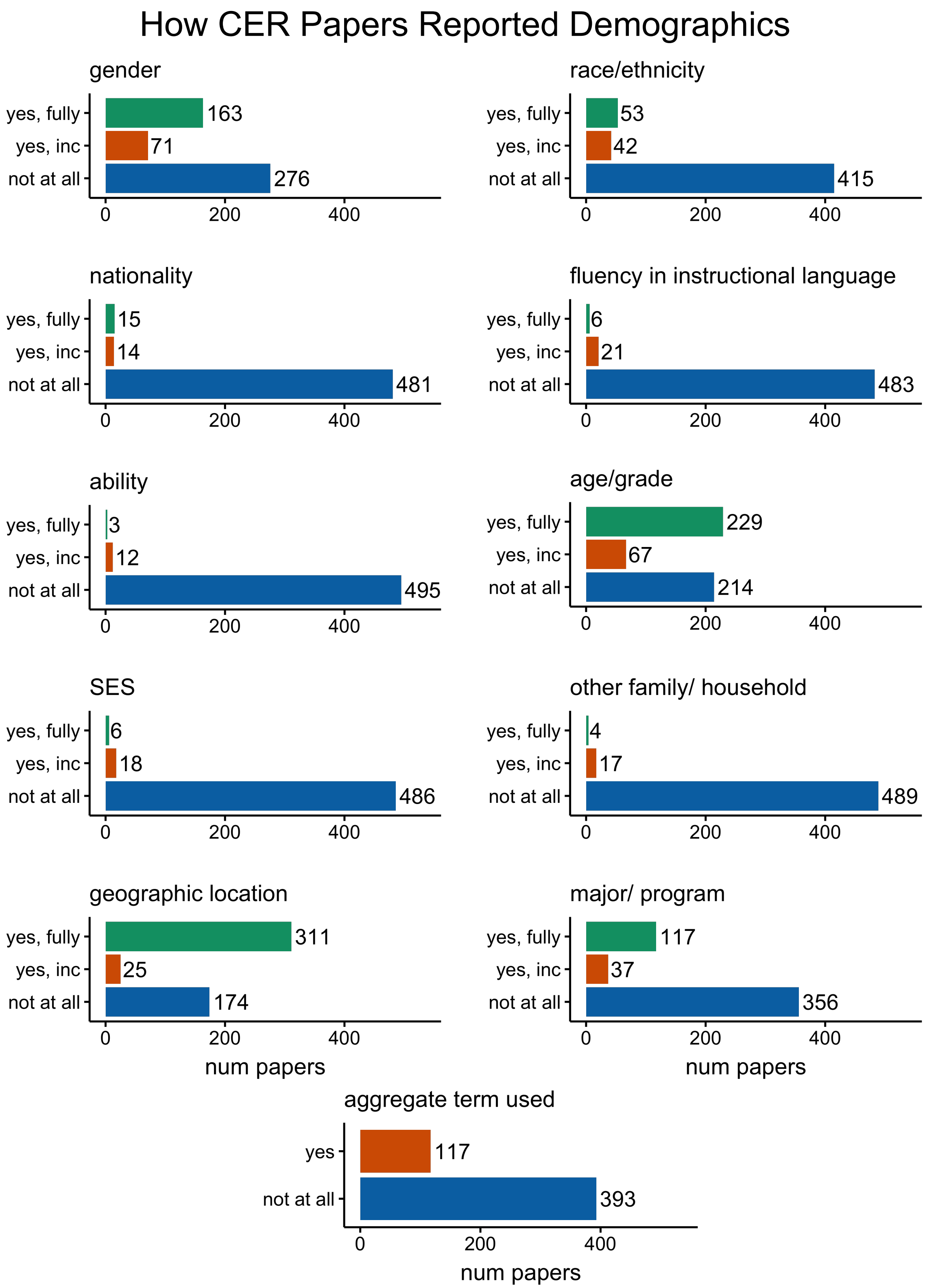

Figure 3: RQ3 Results: How CER papers reported 11 demographics.

erasing non-binary, transgender, gender-questioning, and other types of minority-gender students [174].

About half of the papers analyzed (54%; 276) did not report participants' genders, less prevalent than prior work [53]. Often, the lack of gender information was not explicitly addressed. Papers that reported gender often conflated sex-related terms (male/female) with gender-related terms (man/woman, girl/boy), a pitfall identified in prior work [117].

Within the 32% (163) of papers that fully reported participant gender, categories often reflected binary dichotomies (e.g. boys/girls, male/female). When there was a third category, it was often some form of "did not disclose", rather than a third option. Some papers recognized the existence of genders beyond the binary in the form of an "Other" label, like in *2019-comped-0008*. *2021-toce-0007* disaggregated their "Other" label within the text, including participants who reported as transgender, agender, or another gender not listed. *2018-iticse-0011* explicitly included "transgender" as a category, and *2020-cse-0002* explicitly reported "non-binary" as a gender category for participants. Notably, papers reporting gender beyond the binary largely came from the past five years.

14% (71) of analyzed papers incompletely reported participants' genders. The most common form of incomplete reporting was to only list the proportion of a sample that identified as one gender, relaying either the proportion of women/females/girls, or men/males/boys (e.g. *2019-cserc-0004*). *2016-iticse-0002* only reported the number of female and transgender participants. Some scholars used gender to balance demographics of groups (e.g. *2017-sigcse-0009*) without reporting participants' genders. Incomplete gender reporting implicitly reinforces binary gender norms and contributes to erasure, implying that given information about participants of a single gender, readers can infer the identities of unlabeled participants (typically implied to be the "other" binary gender).

Exemplary papers for gender reporting normalized non-binary genders by allowing participants to self-report and remaining authentic to their chosen labels. For instance, Letaw et al. illustrated the frequency of open-ended responses in students' self-reported genders, and accounted for students who identified as agender or FTM[4] [131]. Menier et al. went beyond simple reporting, explicitly calling for more representation of trans and non-binary learners in CER to counter erasure and avoid perpetuating further marginalization of students [154]. Finally, Register and Ko declined to report gender, but justified their choice because authors considered it irrelevant [184]. This latter approach reflects critical refusal, an approach of refusing participation in labor regimes that reinforce regressive norms [41].

4.3.2 Race and Ethnicity.

Race refers to a group sharing *outward* biological features and some cultural and historical similarities, while ethnicity refers to a group sharing cultural, historical, and familial bonds [22]. Although "race" and "ethnicity" are often conflated, the subtle definition differences indicate connotation differences. The concept of "race" was developed for discriminatory purposes based on physical features, while ethnicity captures the cultural diversity of a population with more accuracy and fewer negative connotations [8, 166, 185, 202]. Worldwide, race and ethnicity have been tied to disparities in education [68, 74, 76, 106, 146, 147]; CER is no exception [82, 143].

Many scholars have argued for the importance of race and ethnicity in CER. Prior work has integrated critical frameworks, such as cultural competence [225] and intersectionality [174, 183, 203]. Others have proposed pedagogies and interventions to address racial and ethnic disparities in computing education [61, 190, 216]. Nonetheless, a literature review of computing outreach activities from 2009-2015 found that 35% (28) of the 80 papers reviewed reported ethnicity of participants [53].

We found a similar trend in our analyzed papers. Most (81%; 415) did not report the race or ethnicity of participants. Only 10% (53) fully reported the race or ethnicity of their participants. For instance, *2015-wipsce-0002* provided a complete breakdown of the race(s) of their participants based on the US census-defined categories. While this paper fully covered their sample, this breakdown is based on US census racial categories, a flawed and controversial tool [12, 209].

Another 8% (42 of 510) incompletely reported the race or ethnicity of their participants, leaving unlabeled participants for assumption. For example, *2017-toce-0001* described their sample as "83% Caucasian". This necessitates that readers make assumptions about the remaining participants, which can rely on hegemonic norms of which racial groups are dominant or marginalized in computing.

Exemplary papers of their participants' race or ethnicity went beyond racial categories and described (or provided proxies for) the different ethnic backgrounds of their participants. Lewis et al. [132] explained the composition of racial categories in their sample. For example, they described "Asian" as "East Asian (e.g., Chinese)," "Southeast Asian (e.g., Cambodian)," "South Asian (e.g., Indian)," or "Other Asian". Ko and Davis [120] supplemented racial categories with the languages spoken at home to illustrate the diversity within the categories.

4.3.3 Nationality.

While nationality *can* mean citizenship, it may also refer to someone's birthplace or residence, the issuing country of a passport, or even someone's ethnicity (e.g., Danish). Nationality also shifts meaning over time, as politics, war, and geography change. CER literature tends to engage with nationality as a context for research rather than an explicit variable. Camp's recently recognized work on reductions in women's participation in CS never explicitly states that all of the cited data and analysis concerns US cultural trends; it is implicit in its citations [39]. As the research community become more global, nationality has surfaced more explicitly through multinational studies (e.g. [193, 201]), which, although many did not collect participant nationality, did strive to include multiple nations.

Nationality was only reported in 6% (29) of our analyzed papers. The papers that reported nationality often had study populations outside of the US, although they often equated country and nationality. For example, *2016-koli-0004* was done in a Finish school, but it was not explicitly stated whether students were Finnish. US-based papers described nationality through citizenship (e.g *2019-sigcse-0020*). However, citizenship does not always equate to nationality [204]. Some simply reported whether the participants were a part of the country (e.g *2019-koli-0002*). Exemplary papers detailed specifically what the participants' nationality was. For example,

[4] Female-to-male, a label some transgender individuals use to describe their gender.

Boateng et al. [35] described all study participants' nationalities as Ghanaian or Ethiopia.

4.3.4 *Fluency in Instructional Language.*

Fluency is an aggregate notion of many distinct skills like reading, speaking, and writing. Research on fluency is often concerned with second language learning, typically English [69]. Prior work has documented impacts of instructional language fluency in education [89] and surfaced subtleties in the nature of fluency in multilingual learners [40]. In CER, language fluency has largely been used to characterize *who* was studied, and not a subject of research itself [129]. Only recently have scholars begun to explore the role of fluency in CER, examining tailored instruction [110], assets of multilingual students [111, 220], and multilingual post-secondary computing education [176, 177].

Language fluency was rarely mentioned in analyzed papers (5.2%; 27). Similar to prior work, language fluency was mostly used to characterize participants, often describing the instructional language when it was not English or the inclusion of English Language Learners (ELL). However, papers were inconsistent in defining or determining ELL status.

Exemplar papers provided nuanced and contextualized descriptions of participants' fluency in the instructional language. For example, Bender et al. provided a rich description of not only the test that measured students' English fluency, but also the limitations of the measurement [26]. Similarly, Laiti et al. used fluency in Finland's indigenous languages to contextualize a study on the "ethno-programming model" [127]. Beyond spoken languages, Ladner et al. detailed thoroughly how teachers used bilingual approaches in Deaf computing education [126].

4.3.5 *Ability.*

Ability is highly complex and multidimensional [19], including diverse motorphysical abilities, developmental constraints on speech and writing, sensory abilities, and cognitive abilities. *Dis*ability is often fraught with stigma, leading to different cultural assumptions and realities about what it means to be disabled. Modern perspectives on ability treat it as a facet of diversity, framing *dis*ability as a byproduct of cultures and infrastructures not designed for this diversity [112]. Recent CER works have examined disabled students' experiences in computing education [124], accessible curriculum and tool development [206], and barriers to data collection on ability [34]. Efforts like AccessCSforALL work towards the inclusion of students with disabilities in the US CS for All movement [125], advancing disability justice goals.

Only 3% (15) of analyzed papers mentioned ability. These papers covered a range of abilities and described them with varying depth, from an aggregate term like "physically disabled" (*2016-iticse-0002*) to naming the specific abilities, like autism or blindness (*2018-sigcse-0003*). Those few papers were often motivated by disability justice and consequently, were exemplary in their careful and thorough discussion of ability. For example, Ludi et al. detailed how students with visual impairments engaged in the development of their programming tool [136]. Ladner et al. was one of the few to study *educators*' abilities, characterizing a professional development for teachers of deaf students [126].

4.3.6 *Age and Grade.*

Age and school grade level are common but imperfect indicators for learners' developmental stages, used in broadly to suggest milestones within moral development [122], psychosocial development [188], and culture [222]. In CER, age itself is rarely a subject of research (e.g. comparing learning of programming across different ages or developmental stages); most studies instead invoke developmental theory to argue for supporting learners at different developmental stages differently (e.g. [75, 135]). Instead, CER often implicitly engages with development through lenses of educational level, often without sensitivity to learner differences within that level, such as how "post-secondary" learners' ages and developmental characteristics can vary widely.

Most papers reported age or grade (58%; 296). Reporting of grades/year of schooling varied worldwide. Most studies were single-site and reported year of school within local norms. Among the 296 papers, 23% (68) relied exclusively on context-specific terms or grade bands, like freshman (year 1 of a 4-year program) in North America, with no further explanation on age (e.g. *2014-koli-0001*). Another 15% (45/296) provided both grade classification and approximate ages, 12% (36/296) only provided an age range (e.g. *2014-wipsce-0004*), and 7% (22/296) included descriptive statistics of age (e.g. *2021-icer-0006*). Notably, 40% (117/296) studied post-secondary learners, but did not provide age. Instead, they defaulted to descriptors like masters students, or only as CS1 students (e.g. *2014-iticse-0009*).

Exemplar papers not only included an age range and grade, but also described them within the study context in language friendly to an international audience. For example, Hogenboom et al. provided the age range of Dutch primary schools and descriptive statistics for ages and grades [100]. Similarly, Von Hausswolff et al. defined the upper secondary school level in Sweden as 16 to 19 years old [221]. While multi-site, international studies were rare, Falkner et al. had a section describing differences in school placements across countries, using age as the common identifier [65].

4.3.7 *Socioeconomic Status (SES) and other Family or Household Information.*

Socioeconomic status refers to students' economic access to resources and corresponding relative societal positions. Poverty has been linked to poor educational outcomes [43], and household-related factors like literacy [46] and post-secondary enrollment [73] have also been linked to different education outcomes, revealing that learning is a socio-cultural phenomenon [159]. CER tends to engage SES through a broadening participation lens. Recent work has examined how to increase educational access for students in poverty [33, 123, 156, 231], how family influences African-American women's persistence in computing [181], and how families shape learning experiences [16, 57].

Socio-economic status and other family or household information was rarely reported in analyzed papers, consistent with a prior review of CER articles on informal learning [151]. Few papers fully reported SES (1%; 6) and family or household information (1%; 4), with an additional 4% (21) and 3% (17) incompletely reporting SES and family or household information, respectively. Many papers reported SES with US-centric terms like "Pell Grants" (financial aid for low-income students in higher education, e.g. *2021-sigcse-0010*) and "free and reduced lunch" (a government program to reduce childhood hunger [98, 164], e.g. *2020-toce-0004*). These terms are not well-known outside the US, limiting interpretability for an international audience. As for family or household information, the most common was "first-generation" (5%; 10/21), referring to students who are first in their family to pursue post-secondary education, but it was often not defined. Other family or household information included parenthood (e.g. *2021-respect-0007*) and computer access (e.g. *2015-icer-0006*).

Exemplar papers provided definitions for regional terms and ample context, improving international interpretability. For example, Salac et al. [192], defined how students with "economic disadvantage" were identified. Alternatively, Beyer et al. used more universal constructs, like parental occupation or education [30]. Lastly, Lyon and Green [140] provided rich descriptions of their participants' SES and family contexts, detailing care-taking and housing responsibilities.

4.3.8 Geographic Location. Invoking location in education broadly engages multiple dimensions of segregation [38]. Characterizations of geography may refer to human density [118], although designations like "urban" can also be a proxy for race and identity, in its origins as a descriptor for Black neighborhoods in American cities [24]. CER scholars have explicitly engaged "urban" as a proxy for culture and class [60] and "rural" through the lens of infrastructure and resources [92]. Multinational studies have also identified disparities in instruction in primary and secondary computing education across different countries and contexts [13, 63, 64, 105].

Over half of analyzed papers (60%; 311) reported geographical location (Figure 3.) Many papers relied heavily on assuming geographic location to be the authors' university, with some using language like "our university" (e.g. *2012-koli-0003*) or "our undergraduates" (e.g. *2016-icer-0001*). Notably, many US-based papers defaulted to regional terms, like "New England" (e.g. *2020-cse-0001*). Using such terms without further context limits understanding in an international audience.

Exemplary papers provided both location and context. For example, Ko and Davis used a neighborhood map to describe the demographics of their population [120]. Others [50, 148, 211] provided historical context about their local education systems, explaining in terms understandable to an international audience.

4.3.9 Major or Program of Study. A major or program of study is the subject of focus in a post-secondary degree. Many scholars have researched the impact of demographics in majors or programs, including students' decision-making process [72] and enrollment in degree programs [11, 25]. Similarly, CER scholars have studied majors or programs of study, with respect to low enrollment of women [14, 42, 218] and Black students [130], student perceptions [94, 152], and enrollment booms [195].

Most analyzed papers (70%; 356) did not report their participants' major or program of study (Figure 3). Only 23% (117) reported major or program fully and another 7% (37) reported incompletely. Of the papers that reported major or program to any extent, a plurality (49%; 76/154) investigated *only* computing-related majors or programs. Majors or programs in analyzed papers spanned post-secondary degrees of different lengths (e.g. *2019-respect-0008*) and levels (e.g. *2014-toce-0003*).

The next most common were papers examining both computing and non-computing majors or programs (27%; 42/154), followed by papers that *only* examined non-computing majors or programs (14%; 22/154). They studied a range of non-computing majors or programs, ranging from theatre (*2015-iticse-0006*) to business (*2020-icer-0004*). However, most of these papers only reported them as "non-computing", without further detail. Given the variety of epistemologies in these non-computing majors or programs, the lack of disaggregation not only reduces clarity, but also implicitly communicates that "non-computing" is a monolith. This monolithic perceptions of "noncomputing" perpetuates hegemonic norms of which epistemologies are valued in computing.

The remaining 14 papers (9%) were unclear what subject the major or program covered. Some mentioned "majors" or "non-majors" without any mentions of subject (e.g. *2012-icer-0002*). This description relies on readers' assumptions and falls back on hegemonic norms of which majors or programs merit study in CER. Others used "STEM" as a descriptor without further detail (e.g. *2019-comped-0009*). However, the inclusion of computing in STEM was inconsistent across analyzed papers, with some separating computing from STEM (e.g. *2021-koli-0003*), while others included computing in STEM (e.g. *2013-sigcse-0001*).

Exemplary reporting of major or program of study provided clear definitions and explanations. Sax et al. listed all 12 majors in their sample [196]. Similarly, Zweben et al. provided categories for each major or program, as well as detailed examples for each category [235].

4.3.10 Use of Aggregate Terms. Aggregate terms are used to describe demographics for various reasons, ranging from pragmatic concerns, like privacy [66], to community solidarity, like the terms "people of color" [214] and "people with disabilities" [29]. However, aggregate terms can also obscure diverse identities and experiences within a community. For example, the term "people of color" includes cultural origins in multiple continents [7] and the term "people with disabilities" spans various forms of disability [29, 103].

While aggregate terms abound in CER, some have critiqued them, citing their denial of explicit personhood, placement of blame on individuals rather than systems, and obfuscation of differences within the groups [224, 227]. We found that 23% (118) of the analyzed papers used an aggregate term to describe a demographic. Further, all demographic characteristics were described using an aggregate term. Aggregate terms were most common in characterizing race or ethnicity (50%; 59/118), major (23%; 27/118), and gender (21%; 25/118). Terms like "underrepresented" and "diverse" were frequently used for race or ethnicity, gender, or both. Terms like "non-computing", and "non-STEM" were frequently used for participants' major or program of study. In addition to demographics, aggregate terms like "at-risk"(*2015-koli-0003*) were used for academic performance but with inconsistent reporting of how it was determined.

Most papers (68%; 81/118) using an aggregate term did not define or disaggregate them. *2015-wipsce-0001* analyzed differences between ethnic groups without stating what those groups were. Further, 14% (16 papers) were unclear *which* demographics the aggregate terms referred to. For example, to describe their participants, *2017-icer-0001* used "homogeneous", while *2019-cserc-0005* used "heterogeneous". However, neither detailed *how* their participants were homogeneous or heterogeneous. Not only does this ambiguity impact the clarity of a paper, they also require readers to assume their meanings, which can implicitly perpetuate norms of dominant and marginalized groups in computing.

Exemplary reporting of aggregate terms included clear definitions or disaggregations. Several [18, 31, 96, 196, 227, 232] provided a definition and a breakdown of the aggregate term "underrepresented" or "diverse". Their definitions varied, ranging from only referring to race or ethnicity [18, 196], to including gender [31, 96, 232], ability [31, 96], sexual orientation [31], people from low-income backgrounds, and multilingual learners [96].

4.4 RQ4: How CER Papers *Used* Demographics

We categorized demographic usage based on the codes in Table 5. Codes were not mutually exclusive, with 44% (222) of papers having one, 28% (141) having two, 13% (67) having 3, and 16% (80) having 4 or more.

4.4.1 Motivation. Among analyzed papers, 30% (152) used demographics as motivation (Figure 4). If papers were motivated by demographics, they almost always provided the corresponding demographics or conducted analysis with them. For instance, Wong motivated their study with youth technology exposure, fully reported the age of their participants, and conducted analysis based on participant age [229].

4.4.2 Description & Contextualization. Demographics were most commonly used for description and contextualization, accounting for 82% (420) and 52% (268) of papers reviewed, respectively. For example, Theodoropoulos et al. [211] both *described* their participants' ages, genders, and geographic location, and *contextualized* the Greek educational system with the Darmstadt model [104]. The norm of using demographics for description and contextualization improves understanding for readers.

4.4.3 Analysis. Among analyzed papers, 23% (118) used demographics in their analysis. For example, Hodari et al. grounded their qualitative analysis in demographics [97], and Hancock et al. explicitly stated their assumptions when using demographics for quantitative analysis [90]. While most papers using demographics for analysis were motivated by demographic phenomena, some analyzed demographics without a demographic motivation. Most in this category (e.g. [23, 36, 51, 58, 81, 113, 221]) evaluated how well their intervention worked for participants across demographics, even if it was not explicitly designed for them. This indicates a norm of inclusive evaluation of interventions in CER. The exception is non-binary gender analysis. When non-binary genders were reported, they were often excluded from gender analysis (e.g. *2020-sigcse-0012*, *2021-wipsce-0001*). These exclusions were largely attributed to small sample size, but the explanations for why size justified exclusion varied. Some cited privacy, while others cited assumptions of parametric statistics. In the latter, the choice to drop non-binary gender data for parametric tests, instead of considering non-parametric alternatives, reinforces hegemonic binary gender norms.

4.4.4 Validity. Surprisingly, only 14% (70) considered demographics regarding the validity their study, even though 94% (478) of studies were single-site in largely western, educated, industrialized, rich, and developed (WEIRD) countries. This suggests a rampant WEIRD bias in CER, consistent with similar fields [32, 95, 134]. The few exceptions to this norm include McGee et al., who compared demographics of participants to broader populations and described how differences between sample and population demographics affect validity of findings [150] and Seraj et al. [199], who acknowledged the lack of geographic and socioeconomic diversity as a limitation.

5 DISCUSSION

Our analysis surfaced several community norms around demographics in CER. CER papers study post-secondary learners in formal contexts more often than primary and secondary learners, though this trend is reversed for educators. Many CER papers left some aspect of demographics collection method unclear, obscuring whether the data consensually represented participants' identities in authentic, unbiased ways. Participant self-reporting through custom instruments was the most used method. Most demographics were rarely reported, except for geographic locations, ages/grades, genders, and area of study, similar to prior literature reviews (e.g. [93, 152, 180]). Many CER papers incompletely reported demographics, especially for race and ethnicity. CER papers also used various aggregate demographics, though definitions were inconsistent if given at all. Finally, following prior work's recommendations [93, 151], most CER papers used demographics to describe samples and contextualize studies. CER papers often left out smaller marginalized groups when using demographics for analysis, and only a small proportion of papers explicitly mentioned how participants' demographics affect the validity of findings, consistent with prior reviews [6, 93].

5.1 Limitations

We did not consider CER papers published before 2012 in our investigation, nor did we sample literature from every venue that publishes CER papers. Even within the past decade, shifting sociocultural norms have influenced the ways that CER papers represent participants' backgrounds and identities. We also did not code for all types of demographics and we only coded papers written in English. This was partially due to pragmatic constraints, but also due to our US-centric training and contexts (Section 3.4). We chose to oversample for smaller and newer venues within our stratified random sampling method, and the amount of publications from each venue that met inclusion criteria often varied from our initial estimate of 70%. The nature of content analysis and our sample precludes us from determining whether papers' reported demographics matched with participants' actual identities, nor whether our coded interpretations fully matched authors' interpretations. Finally, we took a breadth-first approach for this paper. An inherent tradeoff is a lack of depth: We cannot identify precisely how collection and reporting methods might have differed for each type of demographics, and we did conduct fine-grained analyses of trends by venue or by year. We also did not investigate any trends around intersectional identities, since most CER papers engaged with demographics from a single-axis lens [47]. All of these limitations constrain interpretations, though many of them suggest fruitful avenues for future work, especially since CER is a fast-changing field and sociocultural norms around demographics are fluid.

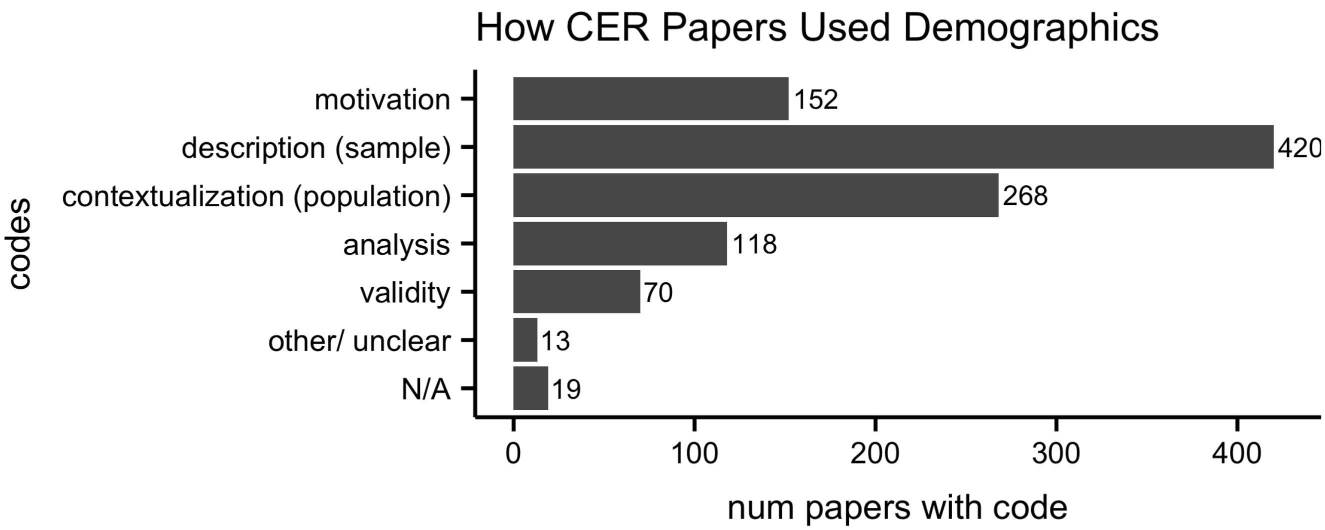

Figure 4: RQ4 Results: How papers used demographics. Total number of codes (1060) exceeds the number of papers analyzed because 288 papers used demographics in multiple ways.

5.2 Considerations

Our goal with this paper was not to prescribe the "best" way to engage with demographics in CER papers, but instead to provide a critical foundation for conversations about the way our field represents students, educators, professionals, and other participants in our studies, without whom most empirical CER work would not be possible. In our field's efforts to broaden participation and support equitable education, Toward this end, we provide several *considerations* that CER researchers should keep in mind when conducting research involving humans. We offer these considerations from the perspective that scientific rigor (a quality that enables the CER community to build off findings for replication, meta-analyses, theory building [9, 93]) and critical reflection are irrevocably intertwined. *Critical reflection is a rigorous practice in itself; rigor contributes to stronger foundations for critical interpretations, and a process cannot be fully rigorous without involving critical reflection.*

5.2.1 When choosing populations, consider who is and isn't there, and why. Many barriers prevent participation in formal post-secondary education, which means that formal post-secondary learners and educators are in privileged positions. Conversely, formal primary and secondary education is mostly compulsory. Instead, barriers to participation are more common in informal experiences, like fees and transportation [56]. Barriers can lead to differences in the demographics of learners across contexts, affecting how representative they are of the general population.

Our results suggest that single-site studies of older (typically post-secondary) learners in formal learning contexts are overstudied in CER relative to other contexts. Conducting research only at a single site can limit the interpretability and applicability of findings, since each context is unique. CER projects that include multiple sites of research and study populations are more likely to generating novel insights that transfer better.

Further, the focus on formal contexts in computing education implicitly privileges specific kinds of learning. The kinds of knowledge that are legitimized in formal (especially post-secondary) education tend to be those steeped in false notions of objectivity, not to mention the centuries-long traditions of gatekeeping and discrimination that have kept people from marginalized backgrounds out of academia. Learners and educators do not exist in a vacuum. As prior work on funds of knowledge [77] and culturally responsive computing [60] have shown, educational outcomes improve when people draw upon their own backgrounds and experiences. To truly broaden participation in computing, CER projects need to center populations beyond "traditional" computing learners.

5.2.2 When collecting demographics, use justified, transparent, and responsible methods. CER researchers should strive towards data collection that respects the humanity of their participants. At minimum, this requires conducting ethical research [10]. It also means weighing the benefits of collecting different demographics against the risks of harm to participants. Researchers should justify their motivations for collecting demographics based on research goals. If there is no justification for collecting particular data, participants should not be forced to disclose their identities. This is especially true if participants receive no benefits from a study or cannot opt out of participation (e.g., a required course).

Transparent demographic collection may involve making instruments for demographics collection publicly accessible, so that others can better build upon the work and better interpret limitations. Initiatives, like CSEdResearch.org, can help researchers archive and use existing instruments [52]. Transparent collection can also involve transparency to participants, by informing them of the reasons behind the collection of their demographics.

Responsible demographic collection allows participants to self-disclose demographics in an authentic way. Collecting demographics through a proxy (like teachers or parents) may lead to inconsistencies with participants' self-conceptions, introducing error to the analysis. Further, since standard classification schemes used for demographics often perpetuate erasure and uphold hegemonic norms (Section 2), researchers should consider these biases if they rely on existing data sets or instruments.

5.2.3 When reporting demographics, recognize biases and make assumptions explicit. The choice of terminology to characterize participants is a value-laden decision. Incomplete reporting of participants' demographics implicitly reinforces hegemonic norms. If readers need to assume characteristics of unmarked participants, they will likely assume that they are part of dominant populations [45]. This can uphold "othering" behavior, dividing and reinforcing demographic groups along existing lines of dominance. Further, critical demographic reporting involves interrogation of the assumptions implicit in the terminology chosen to classify participants. Researchers should consider the values embedded in their classifications, ensuring that the language they use does not implicitly privilege dominant groups or erase marginalized groups. However, complete reporting should not come at the expense of forced disclosure. Potential harms of identification disproportionately fall upon minoritized groups, and people may not want a demographic catalogued in a persistent, archival document like an academic paper. Researchers should always provide participants a means of opting out of demographic disclosure. The number of participants who opt out can be reported alongside the rest of the demographic categories to support completeness in reporting.

Finally, failing to specify the definitions of terminology also decreases clarity for an international audience. Defining terminology also resists the implicit WEIRD-centricity reflected in analyzed papers and supports the replicability and recoverability of findings.

5.2.4 When using demographics, provide details to support interpretation and engage with broader contextual factors. Demographic labels represent reductions of identity facets and are inherently incomplete representations of a person. When considering how interventions might differentially impact groups, demographics should be considered within broader sociocultural settings [101]. This is especially important when hegemonic norms of collection and reporting erase marginalized populations. Analyses that ignore these norms will provide correspondingly limited understandings.

Power structures vary across contexts. Researchers using demographics should report and engage with local power structures, such as racism, misogyny, classism, casteism, or colorism. To surface implicit power dynamics, researchers should reflect on the nature of privilege in their contexts and its impacts on data collection. Including these details in publications can improve interpretations of findings, better contributing to scientific rigor. Future work should provide sociocultural context to demographics and its effect on the relationships revealed through demographic analysis [191].

5.2.5 As a community, support richer participant characterizations, embracing epistemological pluralism. Since empirical research is conducted by and with humans in social, cultural, and political contexts, it can never be a truly objective activity; CER is no exception. However, the recognition of subjectivity is not inherently negative. Assumptions of objectivity can dehumanize both researchers and participants by taking a "view from nowhere" on data and results [115]. More insights into researcher positionality, reflexivity, and context can provide richer interpretations of findings, and support authentic representations of participants. Several non-positivist research paradigms embrace subjectivity, and to encourage these paradigms' use in CER, the community should embrace epistemological pluralism [116, 162, 215]. This shift may require structural support, like the institution of reviewer training on different epistemologies and clear reviewing guidelines for qualitative and critical work.

These considerations, the CER norms we have reported, and recent work surfacing the politics of software and data, and their impacts [27, 165, 168] are a guide to critically considering (and reconsidering) the foundations of computing education research. We hope these works can help us move from research as a reflection of systems of oppression in computing education, to research as an instrument for liberation.

ACKNOWLEDGMENTS

Thanks to Levin Kim, whose feedback and perspectives helped guide our initial demographic analysis. This material is based upon work supported by the National Science Foundation under Grant No. 1539179, 1703304, 1836813, 2031265, 2100296, 2122950, 2137834, 2137312, DGE-1762114, 2127309, and unrestricted gifts from Microsoft, Adobe, and Google.

REFERENCES

[1] 2021. More about SIGCSE. https://sigcse2021.sigcse.org/info/. Accessed: 2022-3-18.

[2] 2022. NSF: Computer Science for All. https://www.nsf.gov/pubs/2020/nsf20539/nsf20539.htm

[3] 2022. UKRI: Big Data and Computing. https://www.ukri.org/our-work/public-engagement/public-engagement-stfc/engage-with-our-areas-of-science-and-technology-stfc/big-data-and-computing/

[4] ACM. 2012. *The 2012 ACM Computing Classification System - Introduction.* Technical Report. https://www.acm.org/publications/class-2012-intro

[5] ACM Publications Board. 2019. Conflict of Interest Policy for ACM Publications. https://www.acm.org/publications/policies/conflict-of-interest. Accessed: 2022-3-18.

[6] Ahmed Al-Zubidy, Jeffrey C. Carver, Sarah Heckman, and Mark Sherriff. 2016. A (Updated) Review of Empiricism at the SIGCSE Technical Symposium. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (Memphis, Tennessee, USA) *(SIGCSE '16)*. Association for Computing Machinery, New York, NY, USA, 120–125. https://doi.org/10.1145/2839509.2844601

[7] Linda Alcoff, Debra A. Barrath, Susan Clements, Jennifer Clancy, F. James Davis, Abby L. Ferber, Carlos A. Fernandez, Freda Scott Giles, David Theo Goldberg, and Susan R. Graham. 1995. *American Mixed Race: The culture of microdiversity.* Rowman & Littlefield.

[8] American Anthropological Association. 1997. *American Anthropological Association Response to OMB Directive 15: Race and Ethnic Standards for Federal Statistics and Administrative Reporting (September 1997).* Technical Report. American Psychological Association.

[9] American Educational Research Association. 2006. Standards for Reporting on Empirical Social Science Research in AERA Publications: American Educational Research Association. *Educational researcher* 35, 6 (Aug. 2006), 33–40. https://doi.org/10.3102/0013189X035006033

[10] American Educational Research Association. 2011. Code of Ethics. *Educational researcher* 40, 3 (April 2011), 145–156. https://doi.org/10.3102/0013189X11410403

[11] Gregory Anderson, Mariana Alfonso, and Jeffrey Sun. 2006. Rethinking Cooling Out at Public Community Colleges: An examination of fiscal and demographic trends in higher education and the rise of statewide articulation agreements. *Teachers College Record* 108, 3 (2006), 422–451.

[12] Margo Anderson and Stephen E. Fienberg. 2000. Race and Ethnicity and the Controversy over the US Census. *Current Sociology* 48, 3 (2000), 87–110.

[13] Tehreem Anwar, Arturo Jimenez, Arsalan Bin Najeeb, Bishakha Upadhyaya, and Monica M. McGill. 2020. Exploring the Enacted Computing Curriculum in K-12 Schools in South Asia: Bangladesh, Nepal, Pakistan, and Sri Lanka. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*. 79–90.

[14] Glen Archer, Leonard Bohmann, Allison Carter, Christopher Cischke, Linda M. Ott, and Leo Ureel. 2016. Understanding Similarities and Differences in Students across First-year Computing Majors. In *2016 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–8.

[15] Monica Babeş-Vroman, Thuytien N. Nguyen, and Thu D. Nguyen. 2021. Gender Diversity in Computer Science at a Large Public R1 Research University: Reporting on a Self-study. *ACM Transactions on Computing Education* 22, 2 (Nov. 2021), 13:1–13:31. https://doi.org/10.1145/3471572

[16] Rahul Banerjee, Leanne Liu, Kiley Sobel, Caroline Pitt, Kung Jin Lee, Meng Wang, Sijin Chen, Lydia Davison, Jason C. Yip, and Amy J. Ko. 2018. Empowering Families Facing Fnglish Literacy Challenges to Jointly Engage in Computer Programming. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[17] Shaowen Bardzell. 2010. Feminist HCI: Taking Stock and Outlining an Agenda for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1301–1310. https://doi.org/10.1145/1753326.1753521

[18] William Bares, Bill Manaris, and Renée McCauley. 2018. Gender Equity in Computer Science through Computing in the Arts–A Six-year Longitudinal Study. *Computer Science Education* 28, 3 (2018), 191–210.

[19] Colin Barnes and Geof Mercer. 2010. *Exploring Disability*. Polity.

[20] Georges-Louis Baron, Beatrice Drot-Delange, Monique Grandbastien, and Françoise Tort. 2014. Computer Science Education in French Secondary Schools: Historical and Didactical Perspectives. *ACM Trans. Comput. Educ.* 14, 2 (June 2014). https://doi.org/10.1145/2602486

[21] Valerie Barr. 2018. Different Denominators, Different Results: Reanalyzing CS degrees by gender, race, and ethnicity. *ACM Inroads* 9, 3 (2018), 40–47.

[22] Robert Bartlett. 2001. Medieval and Modern Concepts of Race and Ethnicity. *Journal of Medieval and Early Modern Studies* 31, 1 (2001), 39–56.

[23] Debarati Basu, Harinni K. Kumar, Vinod K. Lohani, N. Dwight Barnette, Godmar Back, Dave McPherson, Calvin J. Ribbens, and Paul E. Plassmann. 2020. Integration and Evaluation of Spiral Theory Based Cybersecurity Modules into Core Computer Science and Engineering Courses. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 9–15.

[24] Michael M. Bell. 1992. The Fruit of Difference: The Rural-Urban Continuum as a System of Identity 1. *Rural Sociology* 57, 1 (1992), 65–82.

[25] Christopher T. Belser, M. Shillingford, Andrew P. Daire, Diandra J. Prescod, and Melissa A. Dagley. 2018. Factors Influencing Undergraduate Student Retention in STEM Majors: Career development, math ability, and demographics. *Professional Counselor* 8, 3 (2018), 262–276.

[26] Lisa Bender, Gursimran Walia, Krishna Kambhampaty, Kendall E. Nygard, and Travis E. Nygard. 2012. Social Sensitivity and Classroom Team Projects: An Empirical Investigation. In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education (SIGCSE '12)*. Association for Computing Machinery, New York, NY, USA, 403–408. https://doi.org/10.1145/2157136.2157258

[27] Ruha Benjamin. 2019. Race After Technology: Abolitionist tools for the new Jim code. *Social Forces* (2019).

[28] Luciana Benotti, María Cecilia Martínez, and Fernando Schapachnik. 2014. Engaging High School Students Using Chatbots. In *Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education (ITiCSE '14)*. Association for Computing Machinery, New York, NY, USA, 63–68. https://doi.org/10.1145/2591708.2591728

[29] Patricia Berne, Aurora Levins Morales, David Langstaff, and Sins Invalid. 2018. Ten Principles of Disability Justice. *WSQ: Women's Studies Quarterly* 46, 1 (2018), 227–230.

[30] Sylvia Beyer. 2014. Why Are Women Underrepresented in Computer Science? Gender Differences in Stereotypes, Self-Efficacy, Values, and Interests and Predictors of Future CS Course-Taking and Grades. *Computer Science Education* 24, 2-3 (2014), 153–192. https://doi.org/10.1080/08993408.2014.963363

[31] Michael David Black. 2016. Seven Semesters of Android Game Programming in CS2. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education*. 5–10.

[32] Emmanuel G. Blanchard. 2012. On the WEIRD Nature of ITS/AIED Conferences. In *Intelligent Tutoring Systems*. Springer Berlin Heidelberg, 280–285. https://doi.org/10.1007/978-3-642-30950-2_36

[33] Jennifer M. Blaney. 2020. Broadening Participation in Computing: The role of upward transfer. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 254–260.

[34] Brianna Blaser and Richard E. Ladner. 2020. Why is Data on Disability so Hard to Collect and Understand?. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Vol. 1. IEEE, 1–8.

[35] George Boateng, Victor Wumbor-Apin Kumbol, and Prince Steven Annor. 2019. Keep Calm and Code on Your Phone: A Pilot of SuaCode, an Online Smartphone-Based Coding Course. In *Proceedings of the 8th Computer Science Education Research Conference (CSERC '19)*. Association for Computing Machinery, New York, NY, USA, 9–14. https://doi.org/10.1145/3375258.3375260

[36] Dennis Bouvier, Ellie Lovellette, John Matta, Jing Bai, Jacqueline Chetty, Stan Kurkovsky, and Jia Wan. 2019. Factors Affecting the Adoption of Peer Instruction in Computing Courses. In *Proceedings of the Working Group Reports on Global Computing Education*. 1–25.

[37] Geoffrey C. Bowker and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. MIT Press.

[38] Julia Burdick-Will and John R. Logan. 2017. Schools at the Rural-Urban Boundary: Blurring the divide? *The ANNALS of the American Academy of Political and Social Science* 672, 1 (2017), 185–201.

[39] Tracy Camp. 2002. The Incredible Shrinking Pipeline. *ACM SIGCSE Bulletin* 34, 2 (2002), 129–134.

[40] Francine Chambers. 1997. What Do we Mean by Fluency? *System* 25, 4 (1997), 535–544.

[41] Marika Cifor, Patricia Garcia, TL Cowan, Jasmine Rault, Tonia Sutherland, Anita Say Chen, Jennifer Rode, Anna Lauren Hoffmann, Niloufar Salehi, and Lisa Nakamura. 2019. Feminist Data Manifest-No. https://www.manifestno.com/

[42] J. McGrath Cohoon. 2002. Recruiting and Retaining Women in Undergraduate Computing Majors. *ACM SIGCSE Bulletin* 34, 2 (2002), 48–52.

[43] Raewyn Connell. 1994. Poverty and education. *Harvard educational review* 64, 2 (1994), 125–150.

[44] Christina Convertino. 2020. Nuancing the Discourse of Underrepresentation: a feminist post-structural analysis of gender inequality in computer science education in the US. *Gender and education* 32, 5 (July 2020), 594–607. https://doi.org/10.1080/09540253.2019.1632417

[45] Sasha Costanza-Chock. 2020. *Design Justice: Community-led practices to build the worlds we need*. MIT Press.

[46] Patricia A. Crawford and Vicky Zygouris-Coe. 2006. All in the Family: Connecting home and school with family literacy. *Early Childhood Education Journal* 33, 4 (2006), 261–267.

[47] Kimberlé Crenshaw. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum* 1989, 1 (1989). https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8

[48] Kimberlé Williams Crenshaw. 1994. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. In *The Public Nature of Private Violence*, Martha Albertson Fineman and Rixanne Mykitiuk (Eds.). Routledge, 93–118.

[49] CSTA. 2020. High School Surveys. https://www.csteachers.org/page/high-school-surveys. Accessed: 2022-3-15.

[50] Quintin Cutts, Judy Robertson, Peter Donaldson, and Laurie O'Donnell. 2017. An Evaluation of a Professional Learning Network for Computer Science Teachers. *Computer Science Education* 27, 1 (2017), 30–53. https://doi.org/10.1080/08993408.2017.1315958

[51] Holger Danielsiek, Laura Toma, and Jan Vahrenhold. 2018. An instrument to assess self-efficacy in introductory algorithms courses. *ACM Inroads* 9, 1 (2018), 56–65.

[52] Adrienne Decker and Monica M. McGill. 2019. A Topical Review of Evaluation Instruments for Computing Education. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (Minneapolis, MN, USA) *(SIGCSE '19)*. Association for Computing Machinery, New York, NY, USA, 558–564. https://doi.org/10.1145/3287324.3287393

[53] Adrienne Decker, Monica M. McGill, and Amber Settle. 2016. Towards a Common Framework for Evaluating Computing Outreach Activities. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education (SIGCSE '16)*. Association for Computing Machinery, New York, NY, USA, 627–632. https://doi.org/10.1145/2839509.2844567

[54] Leigh Ann DeLyser, Bryan Mascio, and Kelsey Finkel. 2016. Introducing Student Assessments with Evidence of Validity for NYC's CS4All. In *Proceedings of the 11th Workshop in Primary and Secondary Computing Education (WiPSCE '16)*. Association for Computing Machinery, New York, NY, USA, 17–26. https://doi.org/10.1145/2978249.2978252

[55] Paul Denny, Brett A. Becker, Michelle Craig, Greg Wilson, and Piotr Banaszkiewicz. 2019. Research This! Questions That Computing Educators Most Want Computing Education Researchers to Answer. In *Proceedings of the 2019 ACM Conference on International Computing Education Research (ICER '19)*. Association for Computing Machinery, New York, NY, USA, 259–267. https://doi.org/10.1145/3291279.3339402

[56] Betsy DiSalvo, Cecili Reid, and Parisa Khanipour Roshan. 2014. They Can't Find Us: The search for informal CS education. In *Proceedings of the 45th ACM technical symposium on Computer science education*. 487–492.

[57] Stefania Druga, Jason Yip, Michael Preston, and Devin Dillon. 2021. The 4As: Ask, Adapt, Author, Analyze-AI Literacy Framework for Families. In *Algorithmic Rights and Protections for Children*. PubPub.

[58] Caitlin Duncan and Tim Bell. 2015. A Pilot Computer Science and Programming Course for Primary School Students. In *Proceedings of the Workshop in Primary and Secondary Computing Education*. 39–48.

[59] Ruth Dunn. 2021. *Minority Studies*. LibreTexts.

[60] Ron Eglash, Juan E Gilbert, Valerie Taylor, and Susan R Geier. 2013. Culturally responsive computing in urban, after-school contexts: Two approaches. *Urban Education* 48, 5 (2013), 629–656.

[61] Sheena Erete, Karla Thomas, Denise Nacu, Jessa Dickinson, Naomi Thompson, and Nichole Pinkard. 2021. Applying a Transformative Justice Approach to Encourage the Participation of Black and Latina Girls in Computing. *ACM Transactions on Computing Education (TOCE)* 21, 4 (2021), 1–24.

[62] Barbara Ericson and Mark Guzdial. 2014. Measuring Demographics and Performance in Computer Science Education at a Nationwide Scale using AP CS Data. In *Proceedings of the 45th ACM technical symposium on Computer science education* (Atlanta, Georgia, USA) *(SIGCSE '14)*. Association for Computing Machinery, New York, NY, USA, 217–222. https://doi.org/10.1145/2538862.2538918

[63] Katrina Falkner, Sue Sentance, Rebecca Vivian, Sarah Barksdale, Leonard Busuttil, Elizabeth Cole, Christine Liebe, Francesco Maiorana, Monica M. McGill, and Keith Quille. 2019. An International Benchmark Study of K-12 Computer Science Education in schools. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*. 257–258.

[64] Katrina Falkner, Sue Sentance, Rebecca Vivian, Sarah Barksdale, Leonard Busuttil, Elizabeth Cole, Christine Liebe, Francesco Maiorana, Monica M McGill, and Keith Quille. 2019. An International Comparison of K-12 Computer Science Education Intended and Enacted Curricula. In *Proceedings of the 19th Koli Calling International Conference on Computing Education Research*. 1–10.

[65] Katrina Falkner, Sue Sentance, Rebecca Vivian, Sarah Barksdale, Leonard Busuttil, Elizabeth Cole, Christine Liebe, Francesco Maiorana, Monica M. McGill, and Keith Quille. 2019. An International Study Piloting the MEasuring TeacheR Enacted Computing Curriculum (METRECC) Instrument. In *Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education (ITiCSE-WGR '19)*. Association for Computing Machinery, New York, NY, USA, 111–142. https://doi.org/10.1145/3344429.3372505

[66] Nina H. Fefferman, Eileen A O'Neil, and Elena N Naumova. 2005. Confidentiality and Confidence: Is data aggregation a means to achieve both? *Journal of public health policy* 26, 4 (2005), 430–449.

[67] Todd Fernandez, Allison Godwin, Jacqueline Doyle, Dina Verdin, Hank Boone, Adam Kirn, Lisa Benson, and Geoff Potvin. 2016. More Comprehensive and Inclusive Approaches to Demographic Data Collection. *School of Engineering Education Graduate Student Series* 60 (2016).

[68] Margot Ford. 2013. Achievement Gaps in Australia: What NAPLAN reveals about education inequality in Australia. *Race Ethnicity and Education* 16, 1 (2013), 80–102.

[69] Pauline Foster. 2020. Oral Fluency in a Second Language: A research agenda for the next ten years. *Language Teaching* 53, 4 (2020), 446–461.

[70] Batya Friedman and Helen Nissenbaum. 1996. Bias in computer systems. *ACM Transactions on Information Systems* 14, 3 (July 1996), 330–347. https://doi.org/10.1145/230538.230561

[71] Nichole M Garcia, Nancy López, and Verónica N Vélez. 2018. QuantCrit: rectifying quantitative methods through critical race theory. *Race Ethnicity and Education* 21, 2 (March 2018), 149–157. https://doi.org/10.1080/13613324.2017.1377675

[72] Veerle Germeijs, Koen Luyckx, Guy Notelaers, Luc Goossens, and Karine Verschueren. 2012. Choosing a major in higher education: Profiles of students' decision-making process. *Contemporary Educational Psychology* 37, 3 (2012), 229–239.

[73] Melinda M Gibbons and Marianne Woodside. 2014. Addressing the needs of first-generation college students: Lessons learned from adults from low-education families. *Journal of College Counseling* 17, 1 (2014), 21–36.

[74] David Gillborn, Nicola Rollock, Paul Warmington, and Sean Demack. 2016. Race, Racism and Education: inequality, resilience and reform in policy. *University of Birmingham* (2016).

[75] Catherine Wilson Gillespie and Sally Beisser. 2001. Developmentally appropriate LOGO computer programming with young children. *Information Technology in Childhood Education Annual* 2001, 1 (2001), 229–244.

[76] Ingrid Gogolin, Sarah McMonagle, and Tanja Salem. 2019. Germany: Systemic, sociocultural and linguistic perspectives on educational inequality. In *The Palgrave handbook of race and ethnic inequalities in education*. Springer, 557–602.

[77] Norma González, Luis C Moll, and Cathy Amanti. 2006. *Funds of knowledge: Theorizing practices in households, communities, and classrooms*. Routledge.

[78] Joanna Goode. 2008. Increasing diversity in k-12 computer science: strategies from the field. In *Proceedings of the 39th SIGCSE technical symposium on Computer science education* (Portland, OR, USA) *(SIGCSE '08)*. Association for Computing Machinery, New York, NY, USA, 362–366. https://doi.org/10.1145/1352135.1352259

[79] Joanna Goode, Max Skorodinsky, Jill Hubbard, and James Hook. 2020. Computer Science for Equity: Teacher Education, Agency, and Statewide Reform. *Frontiers in Education* 4 (2020), 162. https://doi.org/10.3389/feduc.2019.00162

[80] Kevin J. Gucwa and Harry H. Cheng. 2017. Making Robot Challenges with Virtual Robots. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education (SIGCSE '17)*. Association for Computing Machinery, New York, NY, USA, 273–277. https://doi.org/10.1145/3017680.3017700

[81] Mariluz Guenaga, Andoni Eguíluz, Pablo Garaizar, and Juanjo Gibaja. 2021. How do students develop computational thinking? Assessing early programmers in a maze-based online game. *Computer Science Education* 31, 2 (2021), 259–289.

[82] Mark Guzdial. 2020. Talking about race in CS education. *Commun. ACM* 64, 1 (2020), 10–11.

[83] Juergen Habermas. 1985. *The Theory of Communicative Action: Volume 1: Reason and the Rationalization of Society*. Beacon Press.

[84] Juergen Habermas. 1985. *The Theory of Communicative Action: Volume 2: Lifeword and System: A Critique of Functionalist Reason*. Beacon Press.

[85] Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2016. Digital Footprints and Changing Networks During Online Identity Transitions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, San Jose, California, USA, 2895–2907. https://doi.org/10.1145/2858036.2858136

[86] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. 2018. Gender Recognition or Gender Reductionism? The Social Implications of Embedded Gender Recognition Systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, Montreal QC, Canada, 1–13. https://doi.org/10.1145/3173574.3173582

[87] Margaret Hamilton, Andrew Luxton-Reilly, Naomi Augar, Vanea Chiprianov, Eveling Castro Gutierrez, Elizabeth Vidal Duarte, Helen H. Hu, Shoba Ittyipe, Janice L. Pearce, Michael Oudshoorn, and Emma Wong. 2016. Gender Equity in Computing: International Faculty Perceptions and Current Practices. In *Proceedings of the 2016 ITiCSE Working Group Reports (ITiCSE '16)*. Association for Computing Machinery, New York, NY, USA, 81–102. https://doi.org/10.1145/3024906.3024911

[88] David Hammer and Leema K Berland. 2014. Confusing claims for data: A critique of common practices for presenting qualitative research on learning. *Journal of the Learning Sciences* 23, 1 (2014), 37–46.

[89] Hector Hammerly. 1991. *Fluency and Accuracy: Toward Balance in Language Teaching and Learning. Multilingual Matters 73.* ERIC.

[90] Kathleen J Hancock, Rachelle Maranon, Aaron M Montgomery, and Jodi Tims. 2021. Female Scholars in Computer Science: The Role of Family and Other Factors in Achieving Academic Success. In *2021 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. IEEE, 1–8.

[91] Caroline D Hardin. 2021. "Learning from Mistakes Is the Best Thing": Risk and Persistence at Hackathons. In *21st Koli Calling International Conference on Computing Education Research*. Association for Computing Machinery, New York, NY, USA.

[92] Jean Hardy, Susan Wyche, and Tiffany Veinot. 2019. Rural HCI research: Definitions, distinctions, methods, and opportunities. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–33.

[93] Sarah Heckman, Jeffrey C Carver, Mark Sherriff, and Ahmed Al-zubidy. 2021. A Systematic Literature Review of Empiricism and Norms of Reporting in Computing Education Research Literature. *ACM Trans. Comput. Educ.* 22, 1 (Oct. 2021), 1–46. https://doi.org/10.1145/3470652

[94] C Richard G Helps, Robert B Jackson, and Marshall B Romney. 2005. Student expectations of computing majors. In *Proceedings of the 6th conference on Information technology education*. 101–106.

[95] Joseph Henrich, Steven J Heine, and Ara Norenzayan. 2010. Most people are not WEIRD. *Nature* 466, 7302 (July 2010), 29. https://doi.org/10.1038/466029a

[96] Kathryn Hill and Cheri Fancsali. 2021. Bridging Professional Development to Practice: Using School Support Visits to Build Teacher Confidence in Delivering Equitable CS Instruction. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 725–731.

[97] Apriel K Hodari, Maria Ong, Lily T Ko, and Janet Smith. 2015. Enabling courage: Agentic strategies of women of color in computing. In *2015 Research in Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. IEEE, 1–7.

[98] Lee Hoffman. 2012. Free and Reduced-Price Lunch Eligibility Data in Ed" Facts": A White Paper on Current Status and Potential Changes. *Office of Planning, Evaluation and Policy Development, US Department of Education* (2012).

[99] Anna Lauren Hoffmann. 2021. Terms of inclusion: Data, discourse, violence. *New Media & Society* 23, 12 (Dec. 2021), 3539–3556. https://doi.org/10.1177/1461444820958725 Publisher: SAGE Publications.

[100] Sally A. M. Hogenboom, Felienne F. J. Hermans, and Han L. J. Van der Maas. 2021. Computerized Adaptive Assessment of Understanding of Programming Concepts in Primary School Children. *Computer Science Education* 0, 0 (2021), 1–30. https://doi.org/10.1080/08993408.2021.1914461

[101] Paul W Holland. 1986. Statistics and Causal Inference. *J. Amer. Statist. Assoc.* 81, 396 (Dec. 1986), 945–960. https://doi.org/10.1080/01621459.1986.10478354

[102] Hayward Derrick Horton. 1999. Critical Demography: The Paradigm of the Future? *Sociological Forum* 14, 3 (1999), 363–367.

[103] David L Hosking. 2008. Critical disability theory. In *A paper presented at the 4th Biennial Disability Studies Conference at Lancaster University, UK*, Vol. 14. 736.

[104] Peter Hubwieser. 2013. The darmstadt model: a first step towards a research framework for computer science education in schools. In *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives*. Springer, 1–14.

[105] Peter Hubwieser, Michail N Giannakos, Marc Berges, Torsten Brinda, Ira Diethelm, Johannes Magenheim, Yogendra Pal, Jana Jackova, and Egle Jasute. 2015. A global snapshot of computer science education in K-12 schools. In *Proceedings of the 2015 ITiCSE on working group reports*. 65–83.

[106] Ida Hummelstedt, Gunilla Holm, Fritjof Sahlström, and Harriet Zilliacus. 2021. 'Refugees here and Finns there'–categorisations of race, nationality, and gender in a Finnish classroom. *Intercultural Education* 32, 2 (2021), 145–159.

[107] Robert A Hummer. 1996. Black-white differences in health and mortality: A review and conceptual model. *The Sociological Quarterly* 37, 1 (1996), 105–125.

[108] Petri Ihantola, Arto Vihavainen, Alireza Ahadi, Matthew Butler, Jürgen Börstler, Stephen H. Edwards, Essi Isohanni, Ari Korhonen, Andrew Petersen, Kelly Rivers, Miguel Ángel Rubio, Judy Sheard, Bronius Skupas, Jaime Spacco, Claudia Szabo, and Daniel Toll. 2015. Educational Data Mining and Learning Analytics in Programming: Literature Review and Case Studies. In *Proceedings of the 2015 ITiCSE on Working Group Reports (ITICSE-WGR '15)*. Association for Computing Machinery, New York, NY, USA, 41–63. https://doi.org/10.1145/2858796.2858798

[109] Sharin Jacob, Ha Nguyen, Leiny Garcia, Debra Richardson, and Mark Warschauer. 2020. Teaching Computational Thinking to Multilingual Students through Inquiry-based Learning. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Vol. 1. 1–8. https://doi.org/10.1109/RESPECT49803.2020.9272487

[110] Sharin Rawhiya Jacob, Jonathan Montoya, Ha Nguyen, Debra Richardson, and Mark Warschauer. 2022. Examining the What, Why, and How of Multilingual Student Identity Development in Computer Science. *ACM Transactions on Computing Education (TOCE)* (2022).

[111] Sharin Rawhiya Jacob, Sara Vogel, Rose K Pozos, Patricia Ordóñez Franco, and Jean Ryoo. 2021. Leveraging Multilingual Students' Resources for Equitable Computer Science Instruction. In *2021 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. IEEE, 1–2.

[112] Catherine Jampel. 2018. Intersections of disability justice, racial justice and environmental justice. *Environmental Sociology* 4, 1 (2018), 122–135.

[113] Ben Jelen, Julia Dunbar, Susan Monsey, Olivia K Richards, and Katie A Siek. 2019. Utilizing the Affinity Research Group Model in a Summer Research Experience for Undergraduates Program. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*. 990–996.

[114] Britney Johnson, Ben Rydal Shapiro, Betsy DiSalvo, Annabel Rothschild, and Carl DiSalvo. 2021. Exploring Approaches to Data Literacy Through a Critical Race Theory Perspective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21, Article 706)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3411764.3445141

[115] Stephanie T Jones et al. 2021. We tell these stories to survive: Towards abolition in computer science education. *Canadian Journal of Science, Mathematics and Technology Education* 21, 2 (2021), 290–308.

[116] Yasmin Kafai, Chris Proctor, and Debora Lui. 2020. From theory bias to theory dialogue: embracing cognitive, situated, and critical framings of computational thinking in K-12 CS education. *ACM Inroads* 11, 1 (Feb. 2020), 44–53. https://doi.org/10.1145/3381887

[117] Gopinaath Kannabiran. 2011. Themself: Critical analysis of gender in Facebook. In *CHI workshop paper*.

[118] Patricia J Kannapel and Alan J DeYoung. 1999. The rural school problem in 1999: A review and critique of the literature. *Journal of Research in rural Education* 15, 2 (1999), 67–79.

[119] Os Keyes. 2018. The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 88:1–88:22. https://doi.org/10.1145/3274357

[120] Amy J. Ko and Katie Davis. 2017. Computing Mentorship in a Software Boomtown: Relationships to Adolescent Interest and Beliefs. In *Proceedings of the 2017 ACM Conference on International Computing Education Research (ICER '17)*. Association for Computing Machinery, New York, NY, USA, 236–244. https://doi.org/10.1145/3105726.3106177

[121] Amy J Ko, Alannah Oleson, Mara Kirdani-Ryan, Yim Register, Benjamin Xie, Mina Tari, Matthew Davidson, Stefania Druga, and Dastyni Loksa. 2020. It is time for more critical CS education. *Commun. ACM* 63, 11 (2020), 31–33.

[122] Lawrence Kohlberg. 1974. The claim to moral adequacy of a highest stage of moral judgment. *The journal of philosophy* 70, 18 (1974), 630–646.

[123] Harrison Kwik, Benjamin Xie, and Amy J Ko. 2018. Experiences of Computer Science Transfer Students. In *Proceedings of the 2018 ACM Conference on International Computing Education Research* (Espoo, Finland) *(ICER '18)*. ACM Press, 115–123. https://doi.org/10.1145/3230977.3231004

[124] Richard E Ladner, Caitlyn Seim, Ather Sharif, Naba Rizvi, and Abraham Glasser. 2021. Experiences of Computing Students with Disabilities. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 939–940.

[125] Richard E. Ladner, Andreas Stefik, and Brianna Blaser. 2019. Addressing Disability in CS for All. In *2019 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. 1–1. https://doi.org/10.1109/RESPECT46404.2019.8985671

[126] Richard E. Ladner, Andreas Stefik, Jill Naumann, and Erin Peach. 2020. Computer Science Principles for Teachers of Deaf Students. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Vol. 1. 1–4. https://doi.org/10.1109/RESPECT49803.2020.9272432

[127] Outi Laiti. 2016. The Ethnoprogramming Model. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research (Koli Calling '16)*. Association for Computing Machinery, New York, NY, USA, 150–154. https://doi.org/10.1145/2999541.2999545

[128] Justin Langlois. 2020. Recuperating Conflict: Between Critical Generosity and Antagonistic Activation. *Art/Research International: A Transdisciplinary Journal* 5, 1 (2020), 148–157.

[129] Yinchen Lei and Meghan Allen. 2022. English Language Learners in Computer Science Education: A Scoping Review. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education*. 57–63.

[130] Robert W Lent, Frederick G Lopez, Hung-Bin Sheu, and Antonio M Lopez Jr. 2011. Social cognitive predictors of the interests and choices of computing majors: Applicability to underrepresented students. *Journal of Vocational Behavior* 78, 2 (2011), 184–192.

[131] Lara Letaw, Rosalinda Garcia, Heather Garcia, Christopher Perdriau, and Margaret Burnett. 2021. Changing the Online Climate via the Online Students: Effects of Three Curricular Interventions on Online CS Students' Inclusivity. In *Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021)*. Association for Computing Machinery, New York, NY, USA, 42–59. https://doi.org/10.1145/3446871.3469742

[132] Colleen Lewis, Paul Bruno, Jonathan Raygoza, and Julia Wang. 2019. Alignment of goals and perceptions of computing predicts students' sense of belonging in computing. In *Proceedings of the 2019 ACM Conference on International Computing Education Research*. 11–19.

[133] Calvin A. Liang, Sean A. Munson, and Julie A. Kientz. 2021. Embracing Four Tensions in Human-Computer Interaction Research with Marginalized People. *ACM Transactions on Computer-Human Interaction* 28, 2 (April 2021), 14:1–14:47. https://doi.org/10.1145/3443686

[134] 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. https://doi.org/10.1145/3411764.3445488

[135] Raymond Lister. 2016. Toward a developmental epistemology of computer programming. In *Proceedings of the 11th workshop in primary and secondary computing education*. 5–16.

[136] Stephanie Ludi, Debra Bernstein, and Karen Mutch-Jones. 2018. Enhanced Robotics! Improving Building and Programming Learning Experiences for Students with Visual Impairments. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18)*. Association for Computing Machinery, New York, NY, USA, 372–377. https://doi.org/10.1145/3159450.3159501

[137] Stephanie Lunn, Maíra Marques Samary, and Alan Peterfreund. 2021. Where is Computer Science Education Research Happening?. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (Virtual Event, USA) *(SIGCSE '21)*. Association for Computing Machinery, New York, NY, USA, 288–294. https://doi.org/10.1145/3408877.3432375

[138] Stephanie Lunn, Leila Zahedi, Monique Ross, and Matthew Ohland. 2021. Exploration of Intersectionality and Computer Science Demographics: Understanding the Historical Context of Shifts in Participation. *ACM Trans. Comput. Educ.* 21, 2 (March 2021), 1–30. https://doi.org/10.1145/3445985

[139] Amy Luo. 2019. Content Analysis | A Step-by-Step Guide with Examples. https://www.scribbr.com/methodology/content-analysis/. Accessed: 2022-1-20.

[140] Louise Ann Lyon and Emily Green. 2021. Coding Boot Camps: Enabling Women to Enter Computing Professions. *ACM Trans. Comput. Educ.* 21, 2 (Feb. 2021). https://doi.org/10.1145/3440891

[141] Tia C Madkins, Alexis Martin, Jean Ryoo, Kimberly A Scott, Joanna Goode, Allison Scott, and Frieda McAlear. 2019. Culturally Relevant Computer Science Pedagogy: From Theory to Practice. In *2019 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. 1–4. https://doi.org/10.1109/RESPECT46404.2019.8985773

[142] Martin N Marger. 2015. *Race and Ethnic Relations: American and Global Perspectives, 10th Edition*. Cengage.

[143] Jane Margolis. 2017. *Stuck in the Shallow End, updated edition: Education, Race, and Computing*. MIT press.

[144] Jane Margolis and Allan Fisher. 2002. *Unlocking the clubhouse: Women in computing*. MIT press.

[145] Lauren Margulieux, Tuba Ayer Ketenci, and Adrienne Decker. 2019. Review of measurements used in computing education research and suggestions for increasing standardization. *Computer Science Education* 29, 1 (Jan. 2019), 49–78. https://doi.org/10.1080/08993408.2018.1562145

[146] David Mayeda, Alan France, Tepora Pukepuke, Lucy Cowie, and Marilyn Chetty. 2022. Colonial Disparities in Higher Education: Explaining Racial Inequality for Māori Youth in Aotearoa New Zealand. *Social Policy and Society* 21, 1 (2022), 80–92.

[147] Cameron McCarthy. 1990. *Race and Curriculum: Social Inequality and the Theories and Politics of Difference in Contemporary Research on Schooling.* ERIC.

[148] Roger McDermott, Mark Zarb, Mats Daniels, and Ville Isomöttönen. 2017. First Year Computing Students' Perceptions of Authenticity in Assessment. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '17)*. Association for Computing Machinery, New York, NY, USA, 10–15. https://doi.org/10.1145/3059009.3059062

[149] Steven McGee, Randi McGee-Tekula, Jennifer Duck, Lucia Dettori, Ronald I. Greenberg, Andrew M. Rasmussen, Erica Wheeler, and Adam Shelton. 2020. Does a Computer Science Graduation Requirement Contribute to Increased Enrollment in Advanced Computer Science Coursework?. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Vol. 1. 1–2. https://doi.org/10.1109/RESPECT49803.2020.9272493

[150] Steven McGee, Randi McGee-Tekula, Jennifer Duck, Taylor White, Ronald I. Greenberg, Lucia Dettori, Dale F. Reed, Brenda Wilkerson, Don Yanek, Andrew Rasmussen, and Gail Chapman. 2016. Does a Taste of Computing Increase Computer Science Enrollment?. In *2016 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. 1–8. https://doi.org/10.1109/RESPECT.2016.7836165

[151] Monica M McGill, Adrienne Decker, and Zachary Abbott. 2018. Improving Research and Experience Reports of Pre-College Computing Activities: A Gap Analysis. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (Baltimore, Maryland, USA) *(SIGCSE '18)*. Association for Computing Machinery, New York, NY, USA, 964–969. https://doi.org/10.1145/3159450.3159481

[152] Monica M McGill, Adrienne Decker, and Amber Settle. 2016. Undergraduate students' perceptions of the impact of pre-college computing activities on choices of major. *ACM Transactions on Computing Education (TOCE)* 16, 4 (2016), 1–33.

[153] Marlon Mejias, Ketly Jean-Pierre, Legand Burge, and Gloria Washington. 2018. Culturally Relevant CS Pedagogy - Theory and Practice. In *2018 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. 1–5. https://doi.org/10.1109/RESPECT.2018.8491699

[154] Amanda Menier, Rebecca Zarch, and Stacey Sexton. 2021. Broadening Gender in Computing for Transgender and Nonbinary Learners. In *2021 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. 1–5. https://doi.org/10.1109/RESPECT51740.2021.9620612

[155] Craig S Miller and Amber Settle. 2016. Some Trouble with Transparency: An Analysis of Student Errors with Object-oriented Python. In *Proceedings of the 2016 ACM Conference on International Computing Education Research* (Melbourne, VIC, Australia) *(ICER '16)*. Association for Computing Machinery, New York, NY, USA, 133–141. https://doi.org/10.1145/2960310.2960327

[156] Jeffrey Miller, Saty Raghavachary, and Andrew Goodney. 2018. Benefits of exposing K-12 students to computer science through summer camp programs. In *2018 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–5.

[157] Briana B. Morrison, Lijun Ni, and Mark Guzdial. 2012. Adapting the Disciplinary Commons Model for High School Teachers: Improving Recruitment, Creating Community. In *Proceedings of the Ninth Annual International Conference on International Computing Education Research (ICER '12)*. Association for Computing Machinery, New York, NY, USA, 47–54. https://doi.org/10.1145/2361276.2361287

[158] Briana B Morrison, Beth A Quinn, Steven Bradley, Kevin Buffardi, Brian Harrington, Helen H Hu, Maria Kallia, Fiona McNeill, Oluwakemi Ola, Miranda Parker, et al. 2021. Evidence for Teaching Practices that Broaden Participation for Women in Computing. In *Proceedings of the 2021 Working Group Reports on Innovation and Technology in Computer Science Education*. 57–131.

[159] National Academies of Sciences, Engineering, and Medicine (Ed.). 2018. *How People Learn II: Learners, Contexts, and Cultures.* National Academies Press.

[160] National Science Foundation (NSF). 2022. *Program Solicitation: Broadening Participation in Computing (BPC) (nsf21571).* Technical Report. National Science Foundation (NSF). https://www.nsf.gov/pubs/2021/nsf21571/nsf21571.htm

[161] NCWIT. 2015. Survey-in-a-Box: Student Experience of the Major. https://ncwit.org/resource/sem/. Accessed: 2022-3-1.

[162] Greg L Nelson and Amy J Ko. 2018. On Use of Theory in Computing Education Research. In *Proceedings of the 2018 ACM Conference on International Computing Education Research.* ACM.

[163] Kimberly A Neuendorf. 2017. *The Content Analysis Guidebook.* SAGE.

[164] Lisa M Nicholson, Sandy J Slater, Jamie F Chriqui, and Frank Chaloupka. 2014. Validating Adolescent Socioeconomic Status: Comparing School Free or Reduced Price Lunch with Community Measures. *Spatial Demography* 2, 1 (April 2014), 55–65. https://doi.org/10.1007/BF03354904

[165] Safiya Umoja Noble. 2018. *Algorithms of oppression.* New York University Press.

[166] Anthony Christian Ocampo. 2016. The Latinos of Asia. In *The Latinos of Asia*. Stanford University Press.

[167] Michael Omi and Howard Winant. 2014. *Racial formation in the United States*. Routledge.

[168] Cathy O'neil. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy.* Broadway Books.

[169] Zacharoula Papamitsiou, Michail Giannakos, Simon, and Andrew Luxton-Reilly. 2020. Computing Education Research Landscape through an Analysis of Keywords. In *Proceedings of the 2020 ACM Conference on International Computing Education Research* (Virtual Event, New Zealand) *(ICER '20)*. Association for Computing Machinery, New York, NY, USA, 102–112. https://doi.org/10.1145/3372782.3406276

[170] Dylan Paré. 2021. A Critical Review and New Directions for Queering Computing and Computing Education. *Oxford Research Encyclopedia of Education* (2021).

[171] Thomas H Park and Susan Wiedenbeck. 2011. Learning web development: challenges at an earlier stage of computing education. In *Proceedings of the seventh international workshop on Computing education research* (Providence, Rhode Island, USA) *(ICER '11)*. Association for Computing Machinery, New York, NY, USA, 125–132. https://doi.org/10.1145/2016911.2016937

[172] Michael Quinn Patton. 2014. *Qualitative research & evaluation methods: Integrating theory and practice.* Sage publications.

[173] Dudley L Poston, Jr. 2019. *Handbook of Population.* Springer International Publishing.

[174] Vahab Pournaghshband and Paola Medel. 2020. Promoting Diversity-Inclusive Computer Science Pedagogies: A Multidimensional Perspective. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '20)*. Association for Computing Machinery, New York, NY, USA, 219–224. https://doi.org/10.1145/3341525.3387360

[175] Heather E Price. 2019. Large-Scale Datasets and Social Justice: Measuring Inequality in Opportunities to Learn. In *Research Methods for Social Justice and Equity in Education*, Kamden K Strunk and Leslie Ann Locke (Eds.). Springer International Publishing, Cham, 203–215. https://doi.org/10.1007/978-3-030-05900-2_17

[176] Adalbert Gerald Soosai Raj, Kasama Ketsuriyonk, Jignesh M Patel, and Richard Halverson. 2017. What Do Students Feel about Learning Programming Using Both English and Their Native Language?. In *2017 International Conference on Learning and Teaching in Computing and Engineering (LaTICE)*. IEEE, 1–8.

[177] Adalbert Gerald Soosai Raj, Hanqi Zhang, Viren Abhyankar, Saswati Mukerjee, Eda Zhang, Jim Williams, Richard Halverson, and Jignesh M Patel. 2019. Impact of bilingual cs education on student learning and engagement in a data structures course. In *Proceedings of the 19th Koli Calling International Conference on Computing Education Research*. 1–10.

[178] Vennila Ramalingam and Susan Wiedenbeck. 1998. Development and Validation of Scores on a Computer Programming Self-Efficacy Scale and Group Analyses of Novice Programmer Self-Efficacy. *Journal of Educational Computing Research* 19, 4 (1998), 367–381.

[179] Justus J Randolph. 2007. *Computer science education research at the crossroads: A methodological review of computer science education research, 2000–2005.* Ph.D. Dissertation. Utah State University, Ann Arbor, United States.

[180] Justus J Randolph. 2007. Findings from" A Methodological Review of the Computer Science Education Research: 2000–2005". *ACM SIGCSE Bulletin* 39, 4 (2007), 130–130.

[181] Yolanda Rankin, Maedeh Agharazidermani, and Jakita Thomas. 2020. The Role of Familial Influences in African American Women's Persistence in Computing. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Vol. 1. 1–8. https://doi.org/10.1109/RESPECT49803.2020.9272503

[182] 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. https://doi.org/10.1145/3363033

[183] Yolanda A. Rankin, Jakita O. Thomas, and Sheena Erete. 2021. Black Women Speak: Examining Power, Privilege, and Identity in CS Education. *ACM Trans. Comput. Educ.* 21, 4 (Oct. 2021). https://doi.org/10.1145/3451344

[184] Yim Register and Amy J. Ko. 2020. Learning Machine Learning with Personal Data Helps Stakeholders Ground Advocacy Arguments in Model Mechanics. In *Proceedings of the 2020 ACM Conference on International Computing Education Research (ICER '20)*. Association for Computing Machinery, New York, NY, USA, 67–78. https://doi.org/10.1145/3372782.3406252

[185] Clara E Rodriguez. 2000. *Changing race: Latinos, the census, and the history of ethnicity in the United States.* Vol. 41. NYU Press.

[186] Jennifer Rosato, Chery Lucarelli, Cassandra Beckworth, and Ralph Morelli. 2017. A Comparison of Online and Hybrid Professional Development for CS Principles Teachers. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '17)*. Association for Computing Machinery, New York, NY, USA, 140–145. https://doi.org/10.1145/3059009.3059060

[187] Karen Rosenblum and Toni-Michelle Travis. 2015. *The Meaning of Difference: American Constructions of Race and Ethnicity, Sex and Gender, Social Class, Sexuality, and Disability.*

[188] Doreen A Rosenthal, Ross M Gurney, and Susan M Moore. 1981. From trust on intimacy: A new inventory for examining Erikson's stages of psychosocial development. *Journal of Youth and Adolescence* 10, 6 (1981), 525–537.

[189] Monique Ross, Zahra Hazari, Gerhard Sonnert, and Philip Sadler. 2020. The Intersection of Being Black and Being a Woman: Examining the Effect of Social Computing Relationships on Computer Science Career Choice. *ACM Trans. Comput. Educ.* 20, 2 (Feb. 2020), 1–15. https://doi.org/10.1145/3377426

[190] Jean J Ryoo, Alicia Morris, and Jane Margolis. 2021. "What Happens to the Raspado man in a Cash-free Society?": Teaching and Learning Socially Responsible Computing. *ACM Transactions on Computing Education (TOCE)* 21, 4 (2021), 1–28.

[191] Rogelio Sáenz and Maria Cristina Morales. 2019. Demography of Race and Ethnicity. In *Handbook of Population*, Dudley L Poston, Jr (Ed.). Springer International Publishing, Cham, 163–207. https://doi.org/10.1007/978-3-030-10910-3_7

[192] Jean Salac, Cathy Thomas, Chloe Butler, and Diana Franklin. 2021. Understanding the Link between Computer Science Instruction and Reading & Math Performance. In *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1.* 408–414.

[193] Kate Sanders, Sally Fincher, Dennis Bouvier, Gary Lewandowski, Briana Morrison, Laurie Murphy, Marian Petre, Brad Richards, Josh Tenenberg, Lynda Thomas, et al. 2005. A multi-institutional, multinational study of programming concepts using card sort data. *Expert Systems* 22, 3 (2005), 121–128.

[194] Kate Sanders, Judy Sheard, Brett A Becker, Anna Eckerdal, Sally Hamouda, and Simon. 2019. Inferential Statistics in Computing Education Research: A Methodological Review. In *Proceedings of the 2019 ACM Conference on International Computing Education Research* (Toronto ON, Canada) *(ICER '19)*. Association for Computing Machinery, New York, NY, USA, 177–185. https://doi.org/10.1145/3291279.3339408

[195] Linda J. Sax, Jennifer M. Blaney, Christina Zavala, and Kaitlin N. S. Newhouse. 2020. Who Takes Intro Computing? Examining the Degree Plans of Introductory Computing Students in Light of Booming Enrollments. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Vol. 1. 1–7. https://doi.org/10.1109/RESPECT49803.2020.9272431

[196] Linda J Sax, Kathleen J Lehman, and Christina Zavala. 2017. Examining the enrollment growth: non-cs majors in CS1 courses. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education.* 513–518.

[197] 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 (CHI '17).* Association for Computing Machinery, New York, NY, USA, 5412–5427. https://doi.org/10.1145/3025453.3025766

[198] Kristin A Searle, Deborah A Fields, Debora A Lui, and Yasmin B Kafai. 2014. Diversifying high school students' views about computing with electronic textiles. In *Proceedings of the tenth annual conference on International computing education research* (Glasgow, Scotland, United Kingdom) *(ICER '14).* Association for Computing Machinery, New York, NY, USA, 75–82. https://doi.org/10.1145/2632320.2632352

[199] Mazyar Seraj, Eva-Sophie Katterfeldt, Kerstin Bub, Serge Autexier, and Rolf Drechsler. 2019. Scratch and Google Blockly: How Girls' Programming Skills and Attitudes are Influenced. In *Proceedings of the 19th Koli Calling International Conference on Computing Education Research.* 1–10.

[200] Sadia Sharmin, Daniel Zingaro, Lisa Zhang, and Clare Brett. 2019. Impact of Open-Ended Assignments on Student Self-Efficacy in CS1. In *Proceedings of the ACM Conference on Global Computing Education (CompEd '19).* Association for Computing Machinery, New York, NY, USA, 215–221. https://doi.org/10.1145/3300115.3309532

[201] Leonardo Silva, Antonio Mendes, Anabela Gomes, Gabriel Fortes, Chan Tong Lam, and Calana Chan. 2021. Exploring the Association between Self-Regulation of Learning and Programming Learning: A Multinational Investigation. In *2021 IEEE Frontiers in Education Conference (FIE).* IEEE, 1–8.

[202] Audrey Smedley and Brian D Smedley. 2005. Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American psychologist* 60, 1 (2005), 16.

[203] Amber Solomon, DeKita Moon, Armisha L Roberts, and Juan E Gilbert. 2018. Not just Black and not just a woman: Black women belonging in computing. In *2018 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT).* IEEE, 1–5.

[204] Peter J Spiro. 1997. Dual nationality and the meaning of citizenship. *Immigr. & Nat'lity L. Rev.* 18 (1997), 491.

[205] Claude Steele. 2011. Stereotype Threat and African-American Student Achievement. In *The Inequality Reader* (2 ed.). Routledge, 276–281. https://doi.org/10.4324/9780429494468-31

[206] Andreas Stefik, Richard E Ladner, William Allee, and Sean Mealin. 2019. Computer science principles for teachers of blind and visually impaired students. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education.* 766–772.

[207] Steve Stemler. 2000. An overview of content analysis. *Practical Assessment, Research, and Evaluation* 7, 1 (2000), 17. https://doi.org/10.7275/z6fm-2e34

[208] Steven E Stemler. 2015. Content analysis. *Emerging trends in the social and behavioral sciences: An Interdisciplinary, Searchable, and Linkable Resource* (2015), 1–14.

[209] Hephzibah V Strmic-Pawl, Brandon A Jackson, and Steve Garner. 2018. Race counts: racial and ethnic data on the US Census and the implications for tracking inequality. *Sociology of Race and Ethnicity* 4, 1 (2018), 1–13.

[210] Burçin Tamer and Jane Stout. 2016. *Recruitment and Retention of Undergraduate Students in Computing: Patterns by Gender and Race/Ethnicity.* Technical Report. Computing Research Association.

[211] Anastasios Theodoropoulos, Angeliki Antoniou, and George Lepouras. 2016. How Do Different Cognitive Styles Affect Learning Programming? Insights from a Game-Based Approach in Greek Schools. *ACM Trans. Comput. Educ.* 17, 1 (Sept. 2016). https://doi.org/10.1145/2940330

[212] Jakita O Thomas, Nicole Joseph, Arian Williams, Chan'tel Crum, and Jamika Burge. 2018. Speaking Truth to Power: Exploring the Intersectional Experiences of Black Women in Computing. In *2018 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT).* 1–8. https://doi.org/10.1109/RESPECT.2018.8491718

[213] Debra Thompson. 2015. What Lies Beneath: Equality and the making of racial classifications. *Social Philosophy & Policy* 31, 2 (2015), 114. Publisher: Cambridge University Press.

[214] Joseph S Tuman. 2009. *Communicating terror: The rhetorical dimensions of terrorism.* Sage Publications.

[215] Sherry Turkle and Seymour Papert. 1992. Epistemological Pluralism and the Revaluation of the Concrete. *Journal of Mathematical Behavior* 11, 1 (1992), 3–33.

[216] Sepehr Vakil. 2018. Ethics, identity, and political vision: Toward a justice-centered approach to equity in computer science education. *Harvard Educational Review* 88, 1 (2018), 26–52.

[217] Bas Van Zadelhoff, Ebrahim Rahimi, and Erik Barendsen. 2021. Principles to Facilitate Design-Based Learning Environments for Programming in Secondary Education While Making Learning Visible in an Authentic Way. In *21st Koli Calling International Conference on Computing Education Research.* Association for Computing Machinery, New York, NY, USA.

[218] Roli Varma. 2010. Why so few women enroll in computing? Gender and ethnic differences in students' perception. *Computer Science Education* 20, 4 (2010), 301–316.

[219] Rebecca Vivian and Katrina Falkner. 2019. Identifying Teachers' Technological Pedagogical Content Knowledge for Computer Science in the Primary Years. In *Proceedings of the 2019 ACM Conference on International Computing Education Research* (Toronto ON, Canada) *(ICER '19).* Association for Computing Machinery, New York, NY, USA, 147–155. https://doi.org/10.1145/3291279.3339410

[220] Sara Vogel, Christopher Hoadley, Laura Ascenzi-Moreno, and Kate Menken. 2019. The role of translanguaging in computational literacies: Documenting middle school bilinguals' practices in computer science integrated units. In *Proceedings of the 50th ACM technical symposium on computer science education.* 1164–1170.

[221] Kristina von Hausswolff, Anna Eckerdal, and Michael Thuné. 2020. Learning to Program Hands-on: A Controlled Study. In *Koli Calling '20: Proceedings of the 20th Koli Calling International Conference on Computing Education Research (Koli Calling '20).* Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3428029.3428058

[222] Lev Semenovich Vygotsky and Michael Cole. 1978. *Mind in society: Development of higher psychological processes.* Harvard university press.

[223] Isabel Wagner. 2016. Gender and Performance in Computer Science. *ACM Trans. Comput. Educ.* 16, 3 (May 2016). https://doi.org/10.1145/2920173

[224] Susan E Walden, Deborah A Trytten, Randa L Shehab, and Cindy E Foor. 2018. Critiquing the" Underrepresented Minorities" Label. In *2018 CoNECD-The Collaborative Network for Engineering and Computing Diversity Conference.*

[225] Alicia Nicki Washington. 2020. When twice as good isn't enough: The case for cultural competence in computing. In *Proceedings of the 51st ACM technical symposium on computer science education.* 213–219.

[226] Max Weber. 1948. *From Max Weber: Essays in Sociology.* Vol. 33. Routledge.

[227] Tiffani L Williams. 2020. Underrepresented minority'considered harmful, racist language. *Commun. ACM* (2020).

[228] Alyssa Friend Wise and David Williamson Shaffer. 2015. Why Theory Matters More than Ever in the Age of Big Data. *Journal of Learning Analytics* 2, 2 (Dec. 2015), 5–13.

[229] Billy Wong. 2016. 'I'm good, but not that good': digitally-skilled young people's identity in computing. *Computer Science Education* 26, 4 (2016), 299–317.

[230] Benjamin Xie, Matt J Davidson, Baker Franke, Emily McLeod, Min Li, and Amy J Ko. 2021. Domain Experts' Interpretations of Assessment Bias in a Scaled, Online Computer Science Curriculum. In *Proceedings of the Eight ACM Conference on Learning @ Scale (L@S 2021, Vol. 29).* ACM. https://doi.org/10.1145/3430895.3460141

[231] Rebecca Zarch, Sarah Dunton, Joshua Childs, and Anne Leftwich. 2020. Through the looking glass: Computer science education and the unintended consequences of broadening participation policy efforts. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Vol. 1. IEEE, 1–4.

[232] Yuanlin Zhang, Jianlan Wang, Fox Bolduc, and William G Murray. 2019. LP based integration of computing and science education in middle schools. In *Proceedings of the ACM Conference on Global Computing Education*. 44–50.

[233] Tukufu Zuberi. 2001. *Thicker Than Blood: How Racial Statistics Lie*. U of Minnesota Press.

[234] Stuart Zweben and Betsy Bizot. 2019. *Taulbee Survey*. Technical Report. Computing Research Association.

[235] Stuart H. Zweben and Elizabeth B. Bizot. 2015. Representation of Women in Postsecondary Computing 1990–2013: Disciplines, Institutional, and Individual Characteristics Matter. In *2015 Research in Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. 1–8. https://doi.org/10.1109/RESPECT.2015.7296493