

A Decade of Demographics in Computing Education Research

A Critical Review of Trends in Collection, Reporting, and Use

Alannah Oleson*
olesona@uw.edu
@OAlannah

Benjamin Xie*
bxie@uw.edu
@benjixie

Jean Salac
salac@uw.edu
@SaladwithaC

Jayne Everson
everjay@uw.edu
@everjay

F. Megumi Kivuva
mk5730@bard.edu
@megumikivuva

Amy J. Ko
ajko@uw.edu
@amyjko



Information School
UNIVERSITY of WASHINGTON



PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING



DESIGN
USE
BUILD

Bard

demographic data is
socially situated and
reinforces or
challenges norms

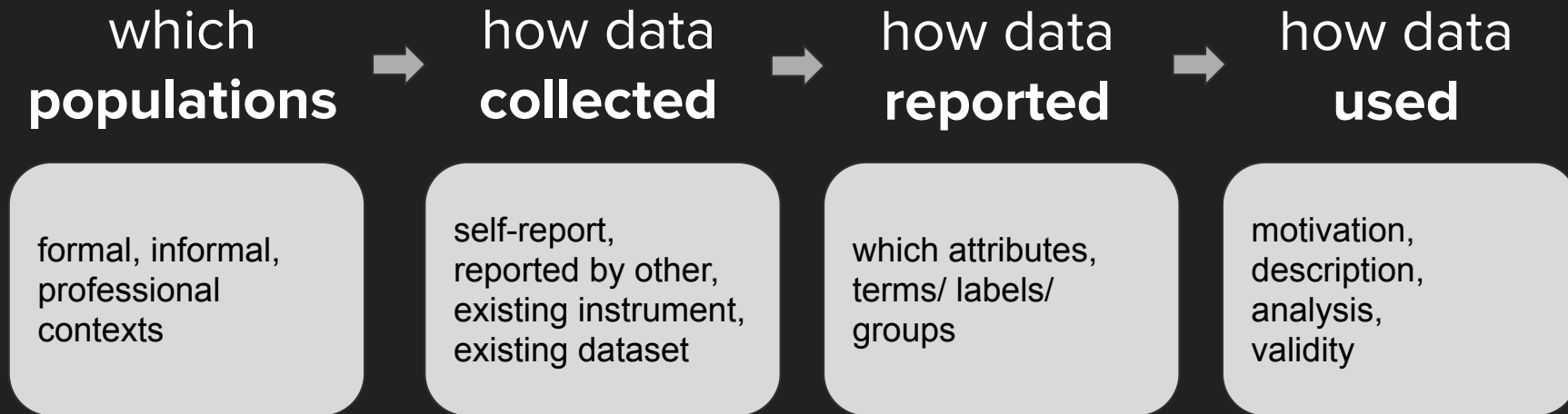
Horton 1999, Zuberi 2001

hegemonic norms:

behaviors that (intentionally or not) perpetuate dominant narratives in CER. They can exclude, harm, or misrepresent people from minoritized groups.

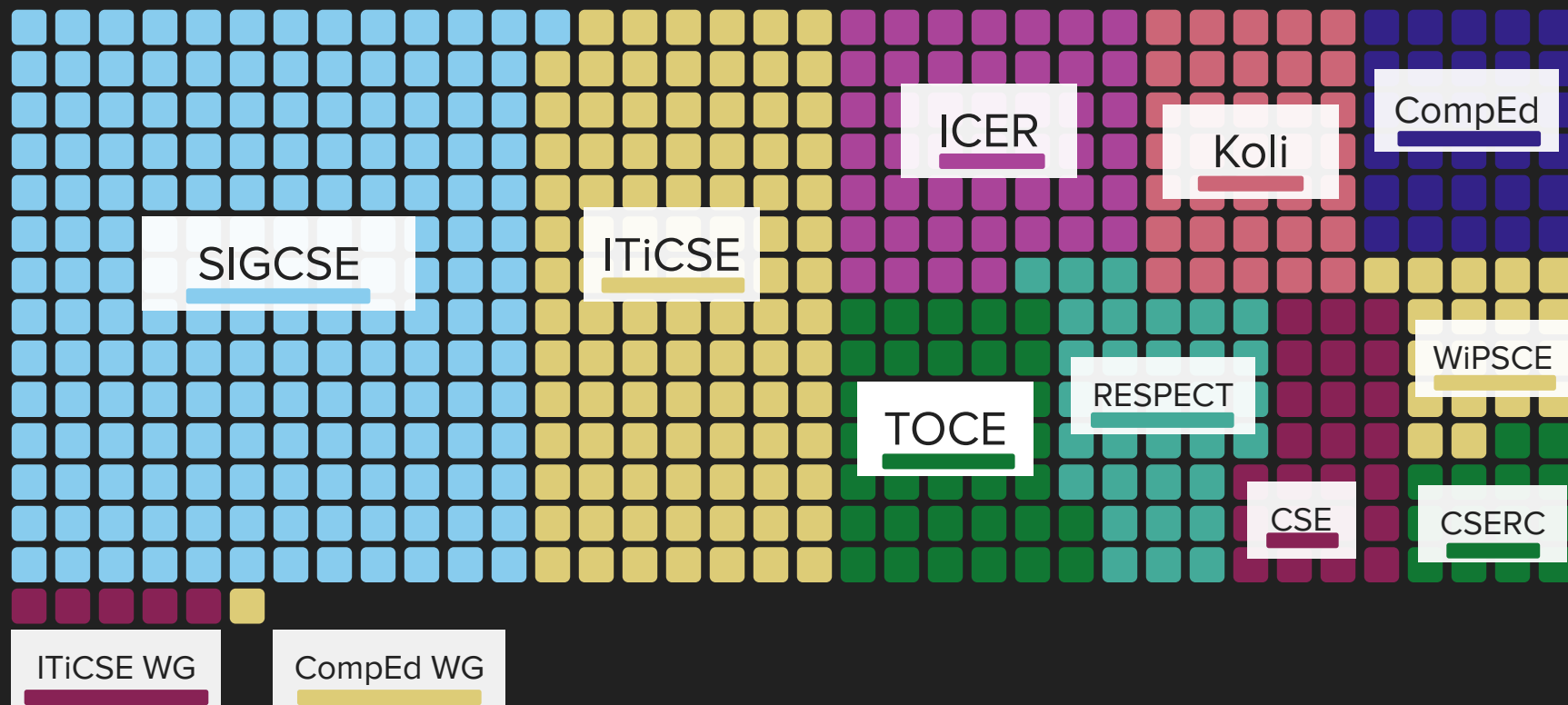
Habermas 1985

unclear how CER wields demographic data



Researchers make **decisions** about demographic data!

content analysis of 510 papers from 12 venues, 2012-2021

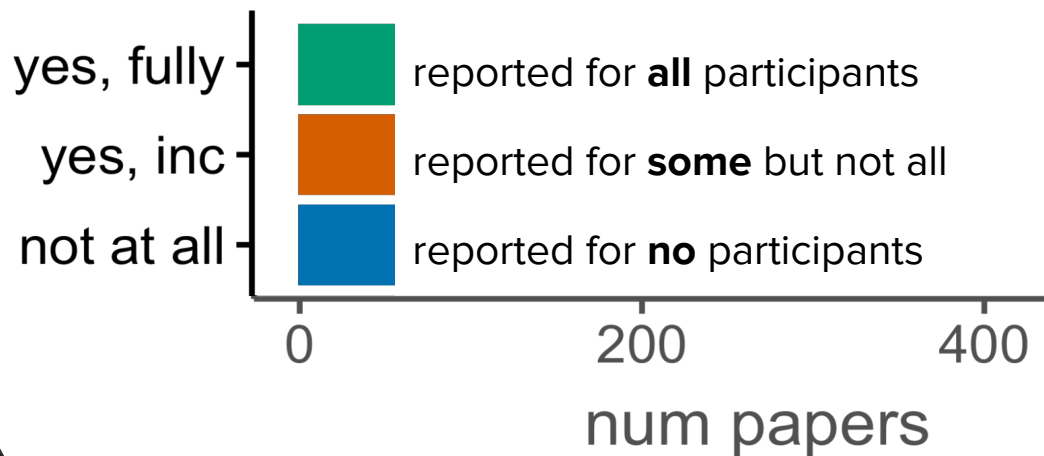




RQ3: What kinds of demographics have been **reported** in CER papers and what kind of language do authors use when reporting?

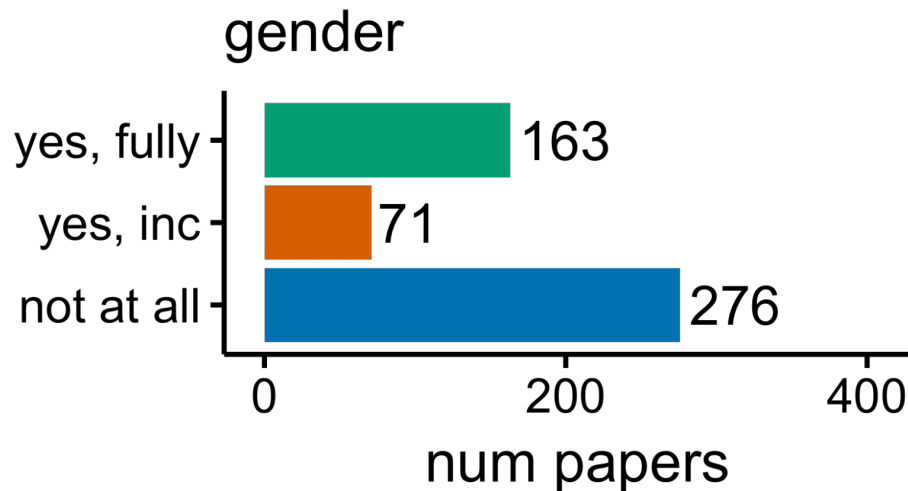
how papers reported **11** demographic attributes

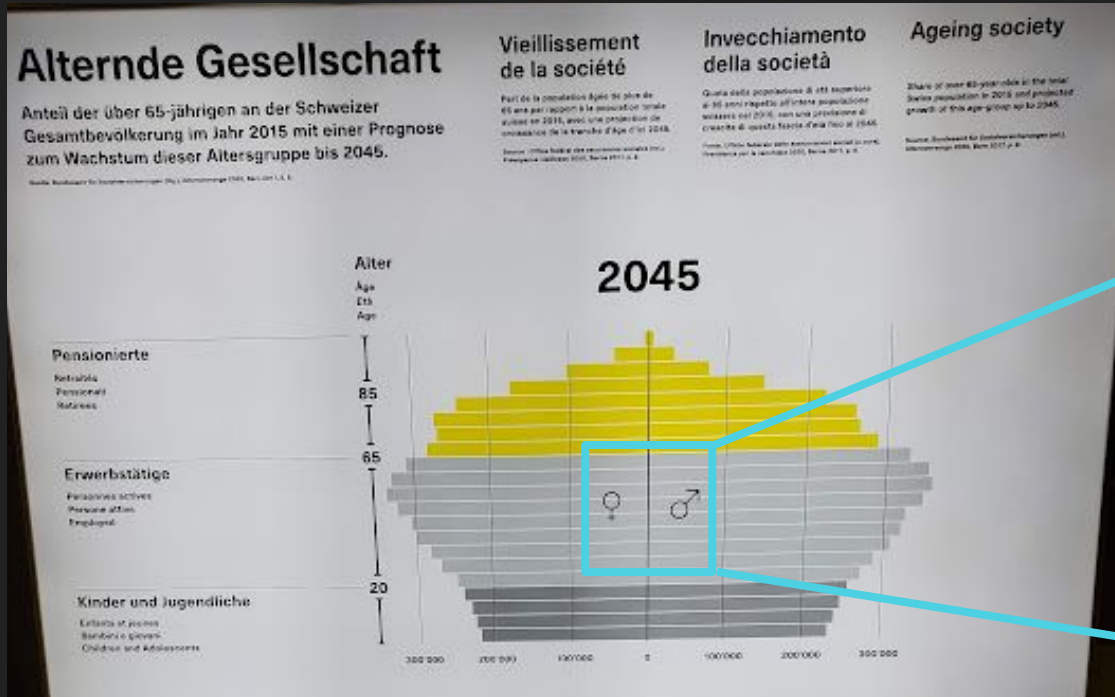
- **gender**
- **race & ethnicity**
- nationality
- fluency in instructional language
- ability
- age/grade
- socio-economic status (SES)
- other family/ household info
- geographic location
- major/ program
- **aggregate term used**



reporting of **gender** partial, assumed dichotomy

- women, men, non-binary, etc
- 3 in 10 papers fully reported
- partial reporting of gender reflected assumption of binary dichotomy
- exemplars challenge norms of dichotomy, required reporting
 - self-report of gender (Letaw et al. ICER 2021)
 - focus on trans and non-binary learners (Menier et al. RESPECT 2021)

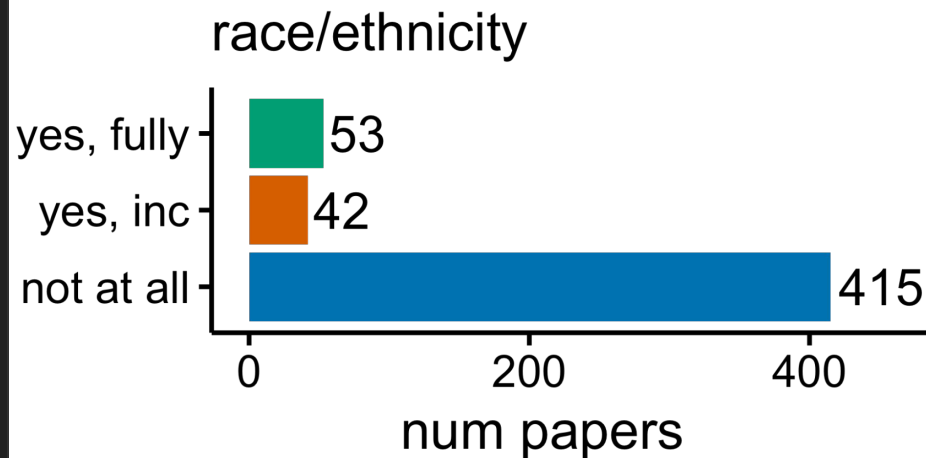




demographics are socially situated

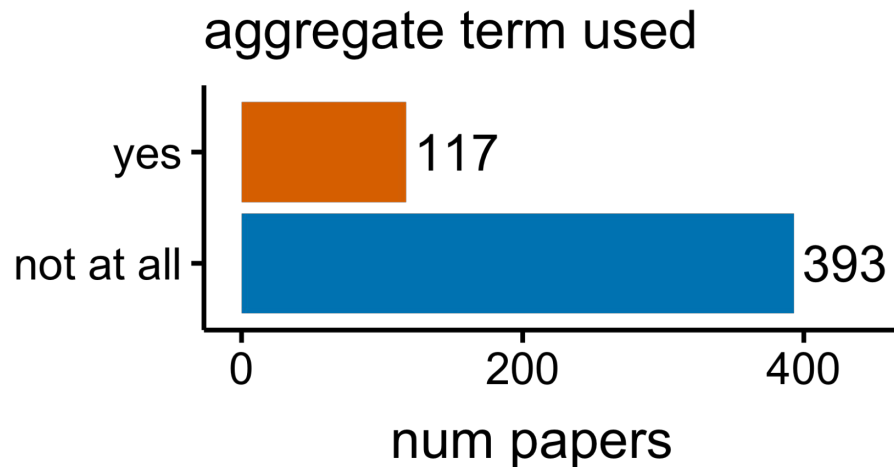
reporting of **race & ethnicity** reflected historical categories

- Black, Indigenous, Hispanic, etc.
- 1 in 10 papers fully reported
- partial reporting reflected dominant norm
 - e.g. "83% Caucasian"
- exemplars went beyond racial categories
 - Explained composition of racial categories (Lewis et al. ICER 2017)
 - Supplemented race with language to illustrate diversity within category (Ko and Davis ICER 2017)



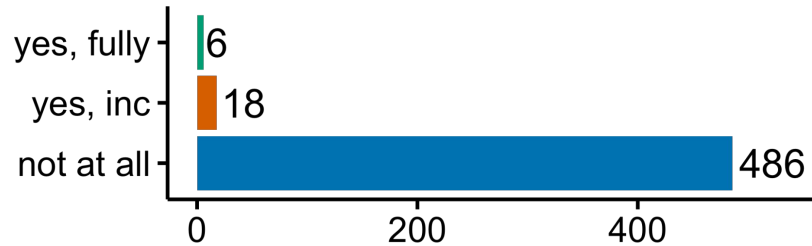
reporting of **aggregate terms** leave ambiguities

- Underrepresented, diverse, at-risk, BIPOC, people w/ disabilities, non-STEM, etc.
- 2 in 10 papers used
- Used to represent all 10 other attributes
 - Most common: ethnicity, major, gender
- Most aggregate terms not defined, leaving readers to make assumptions
- exemplars disaggregated the terms
 - would also like to see justification for use!

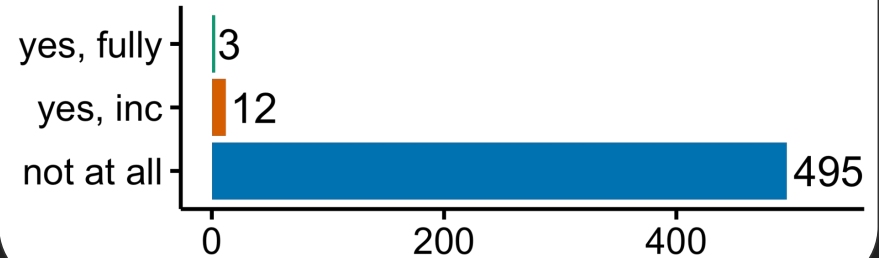


important demographic attributes barely reported

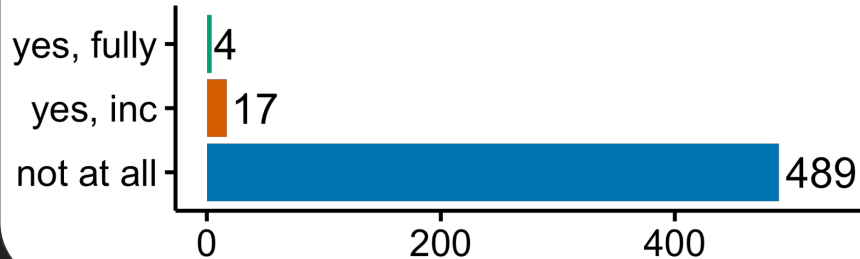
SES



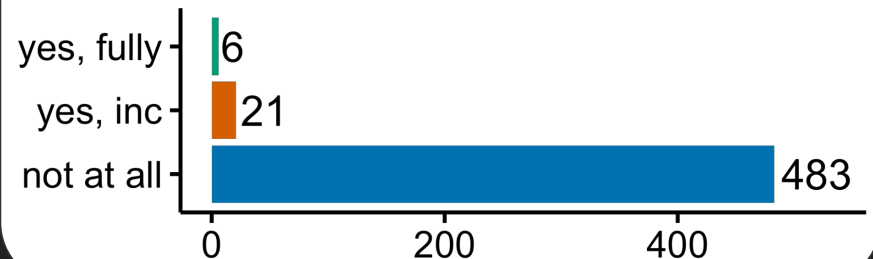
ability



other family/ household



fluency in instructional language



considerations: be reflective and explicit

which
populations



how data
collected



how data
reported



how data
used

when choosing
populations,
consider who is
and isn't there, and
why

justified,
transparent, and
responsible
methods

recognize biases
and make
assumptions
explicit

provide details to
support
interpretation and
engage with
broader context

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

KEYWORDS

demographic data, content analysis, critical demography, literature review

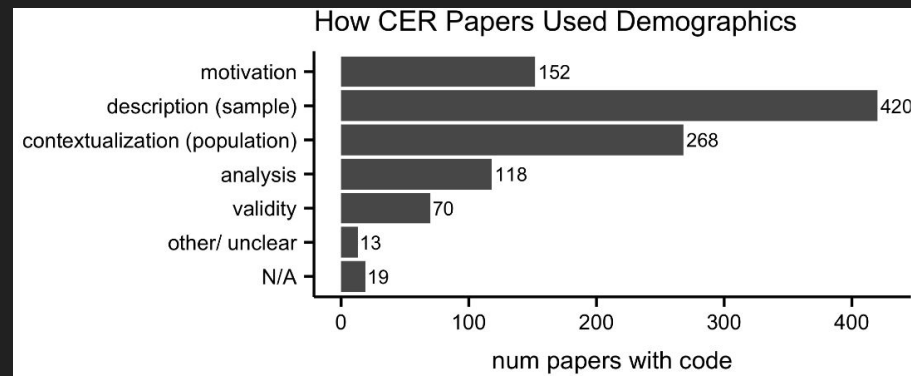
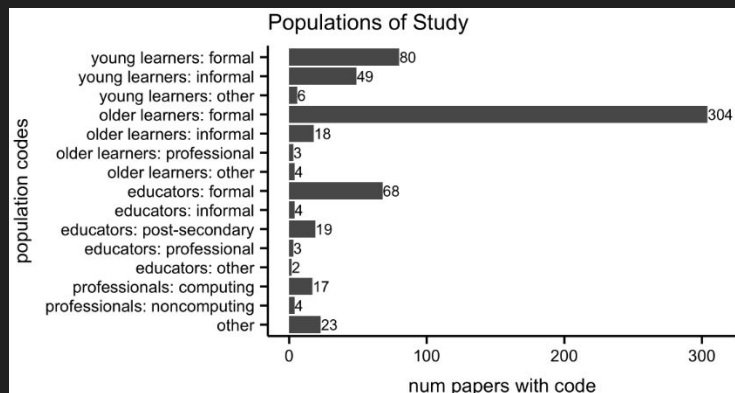
ACM Reference Format:

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 Y1 (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>

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%

read the paper plz



Activity time!

→ tinyurl.com/icer22demo



3 min.

Part 1

Analyze a CER paper to understand how it reports demographics

- Pick a scenario
 - Physical + odd # of people at your table: do **scenario 1** (blue cards)
 - Physical + even # of people: do **scenario 2** (red cards)
 - Virtual folks: Pick your favorite
- Identify *fully reported*, *partially reported*, and *missing* demographics on the padlet

3 min.

Part 2

Reflect on CER's demographic reporting norms and brainstorm how we might do better

- Everyone: What are CER's norms around demographic reporting?
- Everyone: How might we, the CER community, reform these norms?

A Decade of Demographics in Computing Education Research

A Critical Review of Trends in Collection, Reporting, and Use

Alannah Oleson*
olesona@uw.edu
@OAlannah

Benjamin Xie*
bxie@uw.edu
@benjixie

Jean Salac
salac@uw.edu
@SaladwithaC

Jayne Everson
everjay@uw.edu
@everjay

F. Megumi Kivuva
mk5730@bard.edu
@megumikivuva

Amy J. Ko
ajko@uw.edu
@amyjko

full paper & more: benjixie.com/icer22

**which
populations**



**how data
collected**



**how data
reported**



**how data
used**

when choosing
populations,
consider who is
and isn't there, and
why.

justified,
transparent, and
responsible
methods

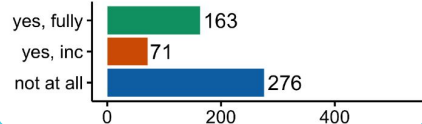
recognize biases
and make
assumptions
explicit.

provide details to
support
interpretation and
engage with
broader context

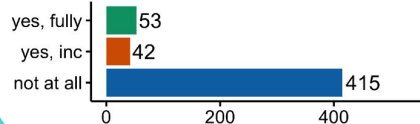
Unused slides

reporting of 11 demographic attributes

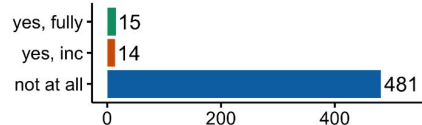
gender



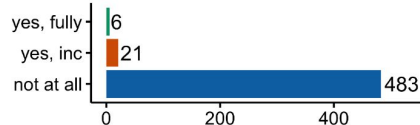
race/ethnicity



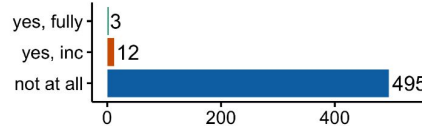
nationality



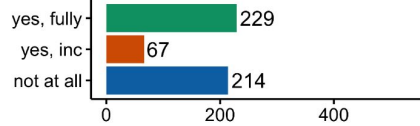
fluency in instructional language



ability

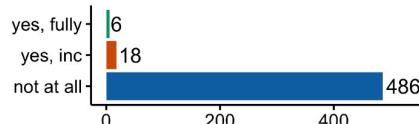


age/grade

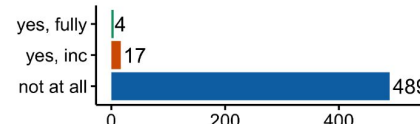


num papers

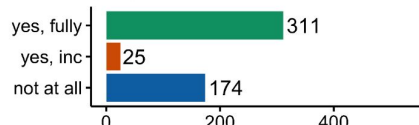
SES



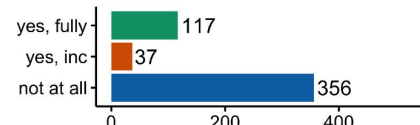
other family/ household



geographic location



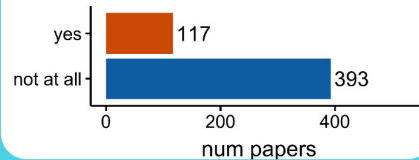
major/ program



num papers

num papers

aggregate term used



num papers

content analysis of 510 papers from 12 venues, 10 yrs

- papers from 12 venues to reflect breadth of CER
 - journals (TOCE, CSE)
 - larger conferences (ICER, SIGCSE, Koli, ITiCSE)
 - newer conferences (RESPECT, CompEd)
 - targeted conferences (CSERC, WIPSCE)
 - working groups (ITiCSE WG, CompEd WG)
 - oversampled from newer, smaller venues
- published 2012-2021
- inclusion criteria: peer-reviewed empirical studies that describe humans
 - Excluded panels, posters (not peer-reviewed)
 - Excluded meta-analyses (not empirical studies of humans)

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.