

# How Administrative Data Collection and Analysis Can Better Reflect Racial and Ethnic Identities

[DOI:10.3102/0091732X20903321]

Samantha Viano    Dominique J. Baker

# Outline

- ▶ Motivation
- ▶ Useful background/definitions
- ▶ Purpose
- ▶ Frameworks
  - ▶ Critical Race Theory
  - ▶ (Traditional) Quantitative Research
  - ▶ QuantCrit
- ▶ Methodology
- ▶ Findings
- ▶ Possible approaches
- ▶ Questions

# Motivation

*Prior research on the use of R/E [race/ethnicity] in the social sciences has often focused on the actual collection and categorization of R/E data (e.g., Denton & Deane, 2010) and how participants are categorized into R/E groups, with less attention paid to how researchers then use those categorizations in their analyses. In this synthesis, we compile the knowledge and insights from this literature to further the field's understanding of how to measure R/E in administrative data and then analyze these data to understand trends and disparities by R/E. (p. 302)*

# Useful background

## Education as a field

- ▶ Low consensus
  - ▶ Many are trained in a discipline
  - ▶ Others trained in education schools
- ▶ Quant/qual divide
- ▶ Practitioner/researcher divide
- ▶ **Upshot:** researchers using a thousand approaches to analyze data collected for a thousand reasons under a thousand assumptions

## Definitions

- ▶ **R/E:** racial and ethnic categories
- ▶ **Administrative data:**

*For the purposes of this study, we use the definition of administrative data used by Figlio et al. (2017). Administrative data sets in education are collected by schools across the K–16 pipeline (a) that include a census of all students (and possibly employees) in that school or institution, (b) that are collected for administrative purposes, and (c) with the school, institution, or their management organizations “owning” the data (though researchers can apply for access). (p. 302)*

# Purpose

*The ultimate purpose of this study is to challenge theorists and methodologists to develop new frameworks that will be more sensitive to complicated R/E identities while also being plausible for those using administrative data sets. (p. 302)*

# Frameworks: Critical Race Theory

*Critical race theory in education positions structural inequality, racism, and White supremacy as inherent aspects of educational system/outcomes, acting as a framework for conceptualizing research and interpreting findings on R/E inequality (Ladson-Billings, 1998; Ladson-Billings & Tate, 1995). (p 304)*

## Three challenges for quantitative use of R/E

1. Claims of neutrality/objectivity
2. Lack of discussion/recognition of power and structural aspects of racism
3. White dominance in the academy

# Frameworks: (Traditional) Quantitative Perspectives

*...education researchers have typically utilized R/E with a lack of attention or understanding that the categories they utilize are superficial and constructed using naive understandings of class and race that are then imbued with deficit-oriented markers of inferiority and superiority. (p 305)*

## Typical statistical concerns

1. Power
2. Precision
3. Parsimony

## Typical outcomes

1. Deficit framing ("achievement gap")
2. Dropping small groups from analysis, or...
3. ...combining into "other" group

# Frameworks: QuantCrit

*A community of scholars has been purposefully attempting to combine critical race theory with quantitative research methods, calling this methodology QuantCrit (Sablan, 2019). The goal of these efforts is to create a space for quantitative research that engages with critical race theory authentically. (p. 305)*

## Process that

1. Recognizes flaws in typical quant processes
2. Notes subjective stance of researcher
3. Takes assets-oriented framing



# Methodology

- ▶ Systematic review
  - ▶ Google Scholar and ProQuest
  - ▶ 2001 to 2019
  - ▶ Peer reviewed
  - ▶ U.S. context
- ▶ 55 articles in final analysis set

# Findings: Measuring R/E

*One clear pattern across the 55 studies included in this research synthesis was the heterogeneity in approaches to measuring R/E. Between changes over time, differences across populations, and methodological choices, measuring R/E was conceptualized in dozens of ways across the 55 studies. (p. 308)*

- ▶ Universal Measures
- ▶ U.S. Census Bureau and Federal Agency Guidance
  - ▶ Race + ethnicity vs mutually exclusive vs select all
  - ▶ Hispanic pan-ethnicity measure is specific challenge
    - ▶ inflated “other”
    - ▶ political concerns with granular approaches
- ▶ Reliability/Validity
  - ▶ R/E identification can change over time...
  - ▶ ...and within context

# Findings: Missing R/E data

*When quantitative education researchers utilize administrative data sets, it is a general expectation that R/E data will, at the very least, be included as covariates regardless of the analytical design. One of the key problems that can arise, however, is missing data on R/E. (p. 318)*

- ▶ Observations often dropped
  - ▶ Problem for small groups
- ▶ Imputation based on
  - ▶ Names
  - ▶ Geocoded addresses
  - ▶ Bayesian Improved Surname Geocoding (BISG)

# Findings: Analysis Incorporating Race/Ethnicity

*Several essays and reviews pointed out a common theme in including R/E in analysis: doing so without nuance, description, or thoughtfulness. While R/E were socially constructed, studies tended to include R/E as covariates with the assumptions that these variables are independent or causal (N. M. Garcia & Mayorga, 2018; James, 2001; Lee, 2009; Ma et al., 2007). (p. 320)*

## Suggestions

- ▶ Do not (w/o explanation):
  - ▶ recode R/E into non-White or “students of color”
  - ▶ blame lack of R/E heterogeneity on lack of power
  - ▶ explain R/E heterogeneity using biological/genetic arguments (just don’t ever this one)
- ▶ Explain relevance of R/E to study
- ▶ Use consistent definitions of R/E

# Possible approaches

## Measurement

- ▶ Disaggregate white group
- ▶ Decide how race is defined (color? self-identification?)
- ▶ Include "most identify/best" follow-up questions
- ▶ Add questions about phenotype

## Missing

- ▶ BISG

## Analysis

- ▶ Incorporate qualitative information
- ▶ Fractional R/E assignment for multiracial individuals
- ▶ Be explicit about modeling choices

# Questions

- ▶ How can we incorporate QuantCrit principals into MRP work?
- ▶ Others?