

Abstract Title Page
Not included in page count.

Title: Learning (Not) to Talk About Race: Investigating What Doctoral Students Learn About Race Variables and Statistical Modeling

Authors and Affiliations: Michael Armijo, University of Pennsylvania; Valerie Lundy-Wagner, New York University; Elizabeth Merrill, University of Pennsylvania

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

SREE's 2012 Spring Conference theme, "Understanding Variation in Treatment Effects", asks many provocative research questions which forward a more nuanced view of schools and students within education evaluation research. Certainly, not all students are alike and not all interventions will have an equal impact on treatment participants, and SREE asks more interesting questions about the systematic variation of impacts. However, beneath these differences in impacts by participant characteristics lurks a culturally sensitive issue that has long been living under the surface in statistical modeling as applied to education research: how socially constructed variables, particularly race variables, are defined, operationalized, and analyzed in predictive models.

Developed for agricultural and other applications, the literature applying these models for social scientific purposes is bereft of considering the cultural constructs that are commonly incorporated into models. This inattention to variables' social meaning is particularly problematic, and unfortunately common, when handling racial variables (Zuberi & Bonilla Silva, 2008). Given that race has such a complicated history and ambiguous meaning, sound research design requires social scientists to describe assumptions about race variables, as well as a theory by which race variables affects the dependent variable of interest. (Barton & Coley, 2010; Hudson, 2003; Kao & Thompson, 2003; Pascarella, 1985; Titus, 2006).

This study investigates how the next generation of educational researchers are being trained around issues of race and statistical modeling. While these issues are messy and complicated, ignoring the complexities of including race variables when training doctoral students can lead students to produce research with ambiguous interpretations and veiled assumptions about race, which ultimately, weakens the state of education research.

There is a large body of literature regarding the meaning of race throughout history and a burgeoning literature of its application to statistical modeling (as noted by Zuberi, 2001). From the beginning, the Black-White binary, historians of science have demonstrated that ideas about racialized groups have changed over time, as have the methods of determining racial membership (Nobles, 2000).

Further complicating this is the Office of Management and Budget's (OMB) adoption of Directive No. 15 that defines race and ethnicity standards to better enforce civil rights laws (Wallman, 1998). The OMB admits that racial categories are sociopolitical constructs for data collection and that the categories are "not to be interpreted as social or biological in nature" (OMB, 1977). The five common categories of American Indian or Alaskan Native, Asian or Pacific Islander, Black, Hispanic and White were recently revised to separate Asians, Hawaiians and Pacific Islanders into distinct groups, as well as allow for bi/multi-raciality in response to political pressure (Prewitt, 2005). Nobles (2000) has documented similar process happening repeatedly since the 1790 decennial census.

The growing literature on race and statistical modeling attempts to bring together the complexities of race and bias with statistical methods. Scholars find that racial variables are often unarticulated-- the variable is not defined in its historical and social context, when included in models, which is cause for concern given that the implicit assumption employed by

SREE Spring 2012 Conference Abstract Template

SREE Spring 2012 Conference Abstract Template 2

SREE Spring 2012 Conference Abstract Template A-2

SREE Spring 2012 Conference Abstract Template B-22

researchers is that race variables are constants with well defined and understood meanings that do not change over time. This anachronistic view of race is wrong, as the social meaning of race changes over time in relation to geographic, demographic, economic, and political factors (Prewitt, 2005; Nobles, 2000; Zuberi, 2001). Further, these unarticulated assumptions can be troublesome because they are liable to racial ideology, reinforcing myths about racial differences.

These theories for groups of people based on their racial affiliation can produce spurious conclusions that masks important relationships and can lead to biased and misleading conclusions. Specifically, use of race as a ‘cause’ for lower standardized achievement scores, high school GPAs, and degree completion rates in higher education is incorrect. When race is used in correlation with or to predict these outcomes, the researcher assumes that race is a well-understood, fixed concept that is both observed and behaves in a uniform manner (Zuberi, 2001). However, qualitative and quantitative research contradicts this treatment of race. For example, Teranishi (2010) presents multiple scenarios of Asian American students, the varied experiences and differential post-secondary outcomes. In a related fashion, work by Massey and colleagues (2007) highlights differences in the experiences and outcomes of Black students whose forefathers were American born slaves and those whose were not.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

This paper asks, how do doctoral students understand the use of race variables in statistical modeling? More specifically, it examines how doctoral students at two universities are trained to define, operationalize, and analyze race variables.

Setting:

Description of the research location.

Two highly-selective graduate schools of education where students are engaged in advanced statistical methods courses.

Population / Participants / Subjects:

Description of the participants in the study: who, how many, key features, or characteristics.

We interviewed students and instructors in addition to conducting a document analysis of their texts and syllabus.

We interviewed six advanced doctoral students from each university who study education and have taken an advanced statistics course, defined as regression analysis, multi-level modeling, or other advanced courses. These students were randomly selected from the pool of doctoral students who have completed their course work.

We interviewed three instructors from each university who teach advanced statistical methods. Again, the process for selection was random.

We collected the syllabus and texts used by all interviewees in their advanced statistics courses.

Intervention / Program / Practice:

Description of the intervention, program, or practice, including details of administration and duration.

This study is a qualitative investigation and does not have an intervention, program, or practice that it is evaluating for an impact analysis.

Research Design:

Description of the research design.

The research design is an exploratory qualitative study that asks how doctoral students understand the use of race variables in statistical modeling. The design incorporates aspects of internal validity through triangulating data using interviews of students and instructors as well as document analysis. External validity is much more difficult given the qualitative nature of this project, but the results of this study should raise questions for those involved with training doctoral students more broadly.

The student interview consists of three parts. The first is a short task, where we ask the student to interpret a standard regression model where race (categorical variables: American Indian, Alaskan Native, Asian, Hawaiians, Pacific Islanders, Black, Hispanic, and White with White as the omitted variable), father’s income, and a categorical geographic variable (urban, suburban, or rural with suburban omitted) predict student test scores. The students were asked to interpret the model, and then asked specifically about the race variable. They were asked to define race, to state possible constructs that the race variable stands for, and to interpret the coefficient and draw conclusions based on their interpretations. In the second component of the interview, the students were shown three types of theoretical justifications for including race in the equations including one based in biological differences, one based in cultural differences, and one based in social-interactions. The students were asked to judge these interpretations of the model and what other variables should be included to test those theories. Lastly, the students were interviewed about the task and their experiences in higher-level statistical classes.

Instructors were shown the task and asked how they believe the students will respond to the prompt. Then the instructors were asked questions regarding how they incorporate discussion about racial variables specifically, or socially constructed variables more broadly, in their courses.

The third data point is a document analysis where we read the texts and syllabi, highlighting key words such as race, socially constructed, defining variables, underlying causal model, theoretical justification, proxy variable, and operationalize, interpreting, and drawing conclusions.

Data Collection and Analysis:

Description of the methods for collecting and analyzing data.

The data collection includes audio-recorded interviews of students and instructors in addition to conducting a document analysis of their texts, syllabus and assignments.

We transcribed the task interactions and interviews into the Atlasti. computer software and then analyzed them by analytic categories (defining, operationalizing, and analysing) and emergent themes.

We wrote analytic memos about each theme and compared them to the document analysis for more general themes relating to the research question, How do doctoral students understand the use of race variables in statistical modeling?

Findings / Results:

Description of the main findings with specific details.

Using responses from the interview questions and short task we uncovered student voices on the prevalence and relevance of investigator perspective in race-related quantitative research personally and in terms of their training. In general, participants misinterpreted the output, employing causal language and incorrectly attributing characteristics to racial group membership. Participants also noted receiving very little, if any mention of the complexity of race as a social construct in their statistical training courses. Neither the syllabus nor the text used in these courses spent adequate space and time exploring the issues associated with using racial variables in statistical modeling.

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

First, researchers must be clear on their definition and concept of race when designing all research, as there are no clear laws, theoretical arguments or consensus on what constitutes race. When racial differences are observed, investigators should examine within-group and between-group variation to explore if racial differences exist or are caused by within group variation. Lack of attention to within group variation masks important relationships and can lead to misleading interpretations, as is the case for Southeast Asians, as noted by Teranishi (2010). Researchers should provide a theoretically grounded rationale in cases where racial groups are excluded from analysis.

Second, there is a long tradition of revealing positionality in qualitative research; a similar standard of should apply to predictive models, especially those including race. While we critique the use of race, we acknowledge that race will continue to be used by state and federal agencies, as well as researchers. However, explicitly describing one's positionality enhances validity research by demonstrating one is aware of a potential problem (Maxwell, 2005) and acknowledging that race is social construct.

Third, acknowledging race and how it is operationalized can lead to better understandings of racialized treatment heterogeneity. Rather than relying on racial classifications in isolation, researchers should emphasize the role of race relations and how they affect student outcomes. For example, when considering differential rates of achievement, we must also obtain data on the racialized nature of academic and social networks that may inhibit historically marginalized college access and persistence (Antonio, 2004). When race is employed as a proxy, researchers should clearly indicate the relationship between race and the variable of interest, as well as implications for policy and practice. Without complementary data and measures to help explain manifestations of race and racism, the quest to reduce achievement gaps in completion and improve outcomes for all students may be a race for which there is no end.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

- Alexander, K.L., Riordan, C., Fennessey, J., & Pallas, A.M. (1982). Social background, academic resources, and college graduation: Recent evidence from the National Longitudinal Survey. *American Journal of Education*, 90(4), 315-333.
- Barton, P.E. & Coley, R.J. (2010). *The Black-White Achievement Gap: When Progress Stopped*. Princeton, NJ: ETS.
- Hudson, L. (2003). Racial/ethnic differences in the path to a postsecondary credential. *Education Statistics Quarterly*, 5(2). Retrieved May 2, 2008, from http://nces.ed.gov/programs/quarterly/Vol_5/5_2/q4_2.asp.
- Kao, G. & Thompson, J.S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, 29, 417-442.
- Massey, D.S., Mooney, M.M., Torres, K.C., & Charles, C.Z. (2007). Black immigrants and Black natives attending selective colleges and universities in the United States. *American Journal of Education*, 113, 243-271.
- Maxwell, J.A. (2005). *Qualitative Research Design: An interactive approach*. Thousand Oaks, CA: Sage Publications.
- Nobles, M. (2000). *Shades of Citizenship: Race and the Census in Modern Politics*. Stanford, CA, Stanford University Press.
- Office of Management and Budget. *Statistical Policy Directive 15*. Washington, DC: Office of Management and Budget; 1977. "Race and Ethnic Standards for Federal Statistics and Administrative Reporting."
- Pascarella, E.T. (1985). Racial differences in factors associated with bachelor's degree completion: A nine-year follow-up. *Research in Higher Education*, 23(4), 351-373.
- Prewitt, K. (2005). Racial classification in America: Where do we go from here? *Daedalus*, 134(1), 5-17
- Teranishi, R.T. (2010). *Asians in the Ivory Tower: Dilemmas of Racial Inequality in American Higher Education*. New York, NY: Teachers College Press.
- Titus, M.A. (2006). Understanding college degree completion of students with low socioeconomic status: The influence of institution financial context. *Research in Higher Education*, 47(4), 371-398.
- Wallman, K.K. (1998). Data on Race and Ethnicity: Revising the Federal Standard. *The American Statistician*, 52(1), 31-33.
- Zuberi, T. (2001). *Thicker than Blood: How Racial Statistics Lie*. Minneapolis, MN: University of Minnesota Press.
- Zuberi, T. (2008). Deracializing social statistics: Problems in the quantification of race, In Zuberi, T. & Bonilla-Silva (Eds.), *White logic, White methods: Racism and methodology*. Lanham, MD: Rowan & Littlefield Publishers.
- Zuberi, T. & Bonilla-Silva, E. (2008). *White logic, White methods: Racism and methodology*. Lanham, MD: Rowan & Littlefield Publishers.

SREE Spring 2012 Conference Abstract Template

SREE Spring 2012 Conference Abstract Template 6

SREE Spring 2012 Conference Abstract Template A-6

SREE Spring 2012 Conference Abstract Template B-66

Appendix B. Tables and Figures
Not included in page count.

<i>SREE Spring 2012 Conference Abstract Template</i>	
<i>SREE Spring 2012 Conference Abstract Template</i>	7
<i>SREE Spring 2012 Conference Abstract Template</i>	A-7
<i>SREE Spring 2012 Conference Abstract Template</i>	B-77