

Racial Disparities in Education: Bridging Theory and Method

Efforts to develop new conceptual and analytical approaches to working with ethnic and racial population data are critical for advancing scientific and public understanding of enduring and emergent patterns of racial and ethnic disparities in an array of consequential social outcomes in the contemporary United States (Downey 2008; Jencks and Phillips 1998; Kao and Thompson 2003; Manza and Uggen 2006; Oliver and Shapiro 2006; Pager and Shepard 2008). This project seeks to overcome an unnecessary barrier to more rapid progress towards this goal: the prevailing disconnect that exists between how race is conceptualized in cutting edge theoretical scholarship, and how race data are operationalized and interpreted in the vast majority of empirical, quantitative studies of racial disparities in social outcomes.

This project draws on constructionist theory to propose a novel way of conceptualizing, modeling and interpreting race and ethnic population data, with the goal of gaining new insight into the social processes that generate and sustain disparities in social outcomes. Of course, the idea that race is a social construction is not news to quantitative researchers. There is broad consensus among the vast majority of social scientists that race and ethnicity are social constructions. Yet with few exceptions, most quantitative studies of racial disparities in social outcomes continue to operationalize, model and interpret race as a fixed, individual characteristic.

This paper integrates constructivist insights into how we deal with population data. The specific insight this paper investigates is that race is not a unidimensional construct, but a multidimensional one. Race as a social relationship is a product of constellation of things: self-perception, other ascription, family ancestry, physical

appearance, and social context. Rather than choose a measure of one of these things as the “best” measure for race, we juxtapose four measures to evaluate how different aspects of race (identity, ascription, skin-color, and family ancestry) may contribute to the apparent “effect of race” on educational outcomes. We articulate the theoretical implications of selecting one measure over another, identifying plausible mechanisms and evaluating alternative measures against existing theories of social processes believed to contribute to educational outcome. Our primary research question is: Do measures of distinct aspects of ‘race’ differ in their associations with educational outcomes, and in theoretically expected ways?

This project utilizes the National Longitudinal Study of Adolescent Health (Add Health) Waves 1-4. Add Health is a school-based, longitudinal study of a nationally representative sample of adolescents in grades 7-12 in the United States during the 1994-95 school year (Harris et al. 2009). The analyses utilize the core longitudinal sample of participants interviewed in Waves 1-4 (N=9,421). The key feature of the Add Health data for our purposes is that it includes multiple measures of race corresponding to alternative conceptions of race, reported by different people and collected multiple times. The Wave 1 and 3 surveys include measures that capture self-identification; the Wave 1, 3 and 4 surveys include measures that capture ascription by others; the Wave 1 parent survey includes measures that capture parental racial self-identification; and Wave 3 includes family ancestry. The richness of the Add Health data in terms of race-related data items provides measures that can be mapped onto conceptually distinct understandings of race.

In this paper, we focus on educational outcomes. Education outcomes are selected as the focus of these analyses because there is a well-developed and analytically sophisticated

literature on racial disparities in educational outcomes and attainment (Jencks and Phillips 1998; Hauser 1993; Roscigno and Ainsworth-Darnell 1999; Gamoran 2001, Hallinan 2001; Downey and Ainsworth-Darnell 2002; Kao and Thompson 2003; Tyson et al. 2005; Morgan 2005; Downey 2008; Herman 2009). For these reasons, educational disparities are a useful case study for an initial determination of the potential value added by the analytical strategy developed in this project for empirical and theoretical understanding of racial disparities in social outcomes.

The innovation in this project is to demonstrate a practical way to *integrate* constructionist understanding of race into quantitative social scientific analysis. More specifically, this study articulates and empirically examines a strategy to conceptualizing race data collected in multiple formats. Based on the constructionist tenant that *race is a social relationship* rather than an individual trait (Barth 1969; Fields 1990; Banks 1996; Jenkins 1997; Cornell and Hartmann 1998), this study proposes a fundamental shift in how researchers think about the problem of inconsistency across measures when dealing with race questions on social surveys.

In most studies, a measure that captures one aspect of the complex social phenomenon of race, typically self-identification, is interpreted as a measure of race per se; the results, in turn, are read as evidence of the significance of race on the outcome. Yet we know from constructionist theory that race is not only a matter of self-identification, it is also a product of other ascription and social context. Race and ethnicity, like other social identities, get formed through the interaction of self-perception and perception by others. As a consequence, 'race' cannot be reduced to one or the other; it is a product of the relationship between them, in iterated social interactions across contexts and time (Fields 1990; Daynes and Lee 2008) The availability of multiple measures of race in many datasets provides an opportunity to glean

additional information about which of the more specific social phenomenon that comprise 'race' contribute more to racial inequality. The conceptual differentiation of distinct measures thus serves as a heuristic tool that provides clues to which dimensions of the relational social construct of race fuel observed racial disparities.

This insight affects how we conceptualize race and ethnicity data collected in social surveys using more than one format, and suggests that the various measures traditionally treated by researchers as substitutable proxies for 'race' do not in fact capture the same underlying 'thing'. Applying this insight implies that the multiple measures available in surveys, such as self-identification, interviewer ascription, skin color, or ancestry are each measures that capture different aspects or dimensions of race as a social relationship. This shifts how we think about the issue of inconsistency across distinct measures; such inconsistency is no longer viewed as a problem to be corrected, but a resource to be exploited, as a means to ascertain more information about how race as a social relationship is operating in a given context, and as an opportunity to investigate if one aspect of the relationship is driving empirical associations with social outcomes more than others.

To answer our primary research question, we estimate multivariate regressions predicting academic achievement and attainment (high school grades, college enrollment, and college completion) using four measures that capture conceptually different dimensions of race (self-identification, other ascription, skin color, and family ancestry). For each outcome we begin by estimating a basic model using one of the four measures of race. The pattern of associations between each measure of race and the educational outcomes estimated in these

models are interpreted as providing suggestive evidence that there is value in theoretically distinguishing the use of these alternative measures.

To determine if the particular dimensions of race used in the analyses are able to be distinguished not only conceptually, but statistically, we next estimate models in which multiple dimensions (measures) of race are included at the same time. This innovative approach has been utilized fruitfully but a handful of times (Telles and Lim 1998; Saperstein 2006; Brown, Hitlin and Elder 2007; Campbell 2009). By setting multiple measures against each other in the same model, we are able to observe how the different dimensions operate simultaneously. One concern may be that the measures are too highly correlated to conduct such analyses, however, in preliminary work with these measures we found that the correlations amongst the four measures vary considerably – the weakest correlation being between self-identification and interviewer ascribed skin color ($r=.42$) and the strongest correlation being between self-identification and interviewer ascription ($r=.89$). These varying correlations provide suggestive evidence that these measures do indeed capture different dimensions of race, and can be meaningfully distinguished from one another statistically.

In preliminary work we find a strong linear association between interviewer ascribed skin color and grades in models that include both self-identification and skin color measures. In other words, holding constant racial self-identification, we find that the darker the skin color, the lower the high school grades. This is evidence of perceived color being a meaningful and distinct indicator of race, and suggests that skin color operates as a meaningful basis of social distinction in addition to racial self-identification. These findings are suggestive of the type of new information that can be gleaned from analyses that use multiple measures to capture distinct aspects of 'race'. The novel analytical strategy developed in this project is not expected

to replace existing approaches to quantitative analysis of racial disparities, rather the aim is to provide additional leverage for weighing alternative hypotheses of underlying social processes that contribute to racial disparities in educational outcomes.