UNIVERSITY OF CALIFORNIA,
IRVINE

Masked Intersectional Inequalities Among Adolescents: Skin Tone Measurement, Skin Color Homophily in Adolescent Friendship Networks, and Skin Color Stratification in Educational Contexts

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Sociology

by

Sara Ivethe Villalta

Dissertation Committee:
Professor David R. Schaefer, Chair
Associate Professor Rachel E. Goldberg
Professor Andrew M. Penner

2022

DEDICATION

To

my mother and father,
whose unwavering love has made this work,
and all good things that come from me, possible

You are your best thing.

Toni Morrison

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACKNOWLEDGMENTS

I would like to express great appreciation and gratitude to my committee chair, Professor David R. Schaefer, and my mentor Rachel E. Goldberg for their immeasurable support throughout my graduate career. They have helped shine a light on an oftentimes dim path. I would also like to thank my committee member, Professor Andrew M. Penner whose belief in my work provided me with the motivation to pursue it. I am also grateful to Professor Marta Tienda, Professor Susan K. Brown, Professor Stan Bailey, Professor Carter T. Butts, and Professor Belinda Campos, whose guidance has helped edify me as a scholar in invaluable ways. Finally, I wish to thank my friends and colleagues Steven Schmidt, Ehsan Farshchi, Martin Jacinto, Amanda Acevedo, Stephanie Pullés, Miles Davison, and Josiah Sweeting for the countless hours spent discussing and refining ideas for the betterment of my research.

VITA

SARA I. VILLALTA

Department of Sociology
University of California, Irvine

ACADEMIC APPOINTMENTS

Postdoctoral Research Associate, Princeton University, beginning Fall 2022, Office of Population Research and Bendheim-Thoman Center for Research on Child Wellbeing

EDUCATION

Ph.D., Sociology, University of California at Irvine, 2022 (expected)
Dissertation: Masked Intersectional Inequalities Among Adolescents:
Skin Tone Measurement, Skin Color Homophily in Adolescent Friendship Networks,
and Skin Color Effects on Different Realms of Adolescent Wellbeing
Committee: David R. Schaefer (chair), Rachel E. Goldberg, Andrew Penner
M.A., Sociology, University of California at Irvine, 2018
B.S., Sociology, with Honors, University of California at Irvine, 2012

RESEARCH INTERESTS

Race and Ethnic, Social Network Analysis, Family and Peer Relationships, Immigration, Education, and Gender

PEER-REVIEWED ARTICLES

Villalta, Sara I., Rachel E. Goldberg, and Marta Tienda. 2022. “Adolescent Partnership Quality and Emotional Health: Insights from a Diary Study.” *Social Forces*.

ARTICLES IN PROGRESS

Villalta, Sara I. “Latino Global and Within-Group Discrimination: How In-Group, Out-Group Dynamics Signal Positionality in the Ethnoracial Hierarchy.” In preparation

Villalta, Sara I., David R. Schaefer, Adriana Umaña-Taylor, and Deborah Rivas-Drake. “Skin Tone Measurement in Heterogenous Populations Using Subject Photos: Examining Predictive Validity Across Multiple Skin Color Scales.” In preparation.

Schaefer, David R., **Sara I. Villalta**, Jonathan D. Ware, Adriana Umaña-Taylor, and Deborah Rivas-Drake. "Birds of Mixed Feathers: Modeling Multiracial Identity and Network Structure." In preparation.

Vezaldenos, Victoria, Bernardette J. Pinetta, Deborah Rivas-Drake, Adriana J. Umaña-Taylor, **Sara I. Villalta**, David R. Schaefer. "The Influence of Skin Color and Discrimination on Biracial Adolescent Ethnic-Racial Identity Saliency." In preparation.

FELLOWSHIPS

Ford Foundation Dissertation Fellowship, National Academy of Sciences	2021-2022
Graduate Division Completion Fellowship, UC Irvine	2021-2022
Graduate Access Preparation Program Fellowship, UC Irvine	2021
Merit Fellowship. School of Social Sciences. UC Irvine	2021
Ford Foundation Dissertation Fellowship, *Alternate and Honorable Mention*	2020-2021
Elliott Family Graduate Fellowship, UC Irvine	2020
Inclusive Excellence Ambassador Fellowship, UC Irvine	2020
Summer Research Fellowship, UC Irvine	2020
Faculty Mentor Program Fellowship, UC Irvine	2019-2020
Ford Foundation Pre-Doctoral Fellowship, *Alternate and Honorable Mention*	2019-2020
Center for Network and Relational Analysis Fellowship, UC Irvine	2017-2022

HONORS AND AWARDS

Outstanding Research Award, Department of Sociology, UC Irvine	2019
DECADE Student Travel Award, UC Irvine	2019
DECADE Student Travel Award, UC Irvine	2018
DECADE Student Travel Award, UC Irvine	2017
Dean's Honors List, UC Riverside, Multiple Quarters	2010-2012
Chancellor's Honors List, UC Riverside	2010-2011

CONFERENCE PRESENTATIONS AND INVITED TALKS

"Examining Reliability and Validity Across Skin Color Scales and Investigating Discrimination by Color and Race," paper presented at the Annual Meeting of the American Sociological Association, Los Angeles — 2022

"Above and Beyond Racial Homophily: How Overlooked Phenotype Stratification Masks Racial Segregation in Adolescent Friendship Networks," paper presented at the Sunbelt International Social Networks Conference, Cairns, Australia — 2022

"The Influence of Friendships on Biracial Adolescent Ethnic-Racial Saliency," (with Victoria Vezaldenos, Bernardette Pinetta, Deborah Rivas-Drake, Adriana — 2022

Umaña-Taylor, and David Schafer), poster presented at the Biennial Meeting of the Society for Research on Adolescence, New Orleans, LA

"Latino Global and Within-Group Discrimination: How In-Group, Out-Group Dynamics Signal Positionality in the Ethnoracial Hierarchy," paper presented at the Annual Meeting of the American Sociological Association, virtual — 2020

"Adolescent Relationship Quality and Emotional Health: Insights from an Intensive Longitudinal Study" (with Marta Tienda and Rachel Goldberg), paper presented at the Annual Meeting of the Population Association of America, Washington D.C., and at the Annual Meeting of the American Sociological Association, New York — 2020 & 2019

"Skin Tone Measurement in Heterogenous Populations Using Subject Photos: Examining Reliability and Construct Validity Across Multiple Skin Color Scales," invited talk at the Center for Network and Relational Analysis, UC Irvine — 2019

"How Did I Get Here? Women and Homelessness," invited talk at the WISEPlace Women's Shelter, Santa Ana, CA — 2019

"Birds of Mixed Feathers: Modeling Multiracial Identity and Network Structure" (with David Schaefer and Jonathan Ware), paper presented at the Sunbelt International Social Networks Conference, Montreal, Canada — 2019

"OC Homelessness Study: Who Are Our Homeless?" invited talk Science Policy Research Group, UC Irvine — 2019

"Social Science Research Methods: Beyond Survey Data," guest lecture for Sociology Research Methods, at California State Polytechnic University, Pomona, CA — 2019

"Differences in Perceptions of Discrimination among Latinos: National-Origin vs. Skin Tone," paper presented at a roundtable at the Annual Meeting of the American Sociological Association, Philadelphia — 2018

"The Costs of Homelessness in Orange County: A Collaborative Study Between UCI and Community Stakeholders" (with David Snow, Rachel Goldberg, and Colin Bernatsky), invited talk at Center for Demographic and Social Analysis, UC Irvine — 2017

Exploratory Examination of the Exclusionary Meanings of Ethnic Identities," paper presented at a roundtable at the Annual Meeting of the Pacific Sociological Association, San Diego, CA — 2012

TEACHING EXPERIENCE AND ADVISING

Instructor of Record

Race & Ethnicity, undergraduate, Summer 2021, UC Irvine
Introduction to Sociology, undergraduate, Summer 2020, UC Irvine

Teaching Assistant

Research Practicum: Race, Color & Phenotype, Winter 2020, Spring 2020, Fall 2020, UC Irvine
Transnational Gangs, Winter 2017, Spring 2018, UC Irvine
Introduction to Sociology, Fall 2015, UC Irvine

Peer Mentor

Sean Avellini, PhD Student, UC Irvine, 2018-2019
Martha Morales, PhD Student, UC Irvine, 2016-2017

Undergraduate Advising

Karen Alondra Villasano, UC Irvine, 2020-2021, now at Indiana University as a PhD student, Department of Sociology
Crista Ureña Hernandez, UC Irvine, 2019-2021, now at UC Irvine as a master's student, Department of Sociology–Demographic and Social Analysis
Elizabeth Cervantes, UC Irvine, 2019-2019, now at UC Irvine as a master's student, Department of Sociology–Demographic and Social Analysis

RESEARCH EXPERIENCE

Graduate Student Researcher for the mDiary Study of Adolescent Relationships, Department of Sociology, Princeton University & UC Irvine — 2017-2022

Project Manager, TIDES Phenotype Study, Department of Sociology, Harvard University, University of Michigan, & UC Irvine — 2018-2021

Graduate Student Researcher for the OC Homelessness Cost Study, Department of Sociology, University of California, Irvine — 2016-2018

Co-Principal Investigator, CACS Wellness Study, Chancellor's Advisory Committee on Staff, UC Irvine — 2017

Graduate Student Researcher, Pine Ridge Reservation Book Project, Department of Sociology, UC Irvine — 2017

Research Assistant for the Intercultural/Interracial Therapy Project, Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Spain — 2012

Translator, West Covina Cultural Identity Study, Department of Sociology, UC Riverside — 2011

RESEARCH SKILLS

Quantitative Methods: Experienced and proficient in survey construction, survey administration, analysis of cross-sectional and intensive longitudinal data, social network analysis, demographic analysis, and population forecasting.

Qualitative Methods: Experienced and proficient in qualitative interviewing (both face-to-face and focus groups). Working knowledge of qualitative data analysis. Fluent in spoken and written Spanish.

PROFFESSIONAL, UNIVERSITY, AND COMMUNITY SERVICE

Palo Verde Resident Council, University Housing, University of California, Irvine — 2017-2022

Ad Hoc Reviewer, *Sociological Perspectives* — 2021

Dean's Fellowship Advisory Committee, School of Social Sciences, University of California, Irvine — 2018-2020

Faculty Representative, Department of Sociology University of California, Irvine — 2018-2019

Summer Undergraduate Research Fellowship Oral Presentation Advisor, University of California, Irvine — 2018-2019

We Are Them: Working Together Across Diverse Backgrounds, Workshop Chair at Santa Ana High School — 2017

Diverse Educational Community and Doctoral Experience, Recruitment Co-Chair, University of California, Irvine — 2016-2018

Organizer, Know-How Sessions, Graduate Division University of California, Irvine — 2016-2018

Delivered presentation, "Know How Sessions: Why Graduate School?," multiple quarters, Graduate Division, University of California, Irvine — 2016-2018

Delivered presentation, "Know How Sessions: Graduate Application 101," multiple quarters, Graduate Division, University of California, Irvine — 2016-2018

Delivered presentation, "Know How Sessions: Writing a Statement of Purpose," multiple quarters, Graduate Division, University of California, Irvine — 2016-2018

Delivered presentation, "Know How Sessions: Funding Your Graduate Education," multiple quarters, Graduate Division, University of California, Irvine	2016-2018
Cohort Representative, Department of Sociology, University of California, Irvine	2016-2017
Welcome Committee Member, Department of Sociology, University of California, Irvine	2015-2016
Co-Founder and President of the Undergraduate Sociological Association, University of California, Riverside	2011-2012

PROFFESSIONAL AFFILIATIONS

American Sociological Association (ASA), member

Population Association of America (PAA), member

International Network for Social Network Analysis (INSNA), member

ABSTRACT OF THE DISSERTATION

Masked Intersectional Inequalities Among Adolescents:

Skin Tone Measurement, Skin Color Homophily in Adolescent Friendship Networks,

and Skin Color Stratification in Educational Contexts

by

Sara Ivethe Villalta

Doctor of Philosophy in Sociology

University of California, Irvine, 2022

Professor David R. Schaefer, Chair

The objective of this dissertation is to uncover how a failure to account for racial appearance, measured as skin color, in studies examining racial stratification within school contexts has the potential to mask inequality among students. The dissertation not only addresses how skin color stratification shapes unequal educational outcomes but also problematizes the issue of how best to capture skin tone data, a measure lacking standardization within the social sciences. Using network data from the Teen Identity Development and Education Study, coupled with data collected using an innovative skin color coding design, the findings of this dissertation underscore how the multidimensionality of race results in overlapping ethnoracial hierarchies that operate differently for certain ethnoracial groups.

INTRODUCTION

The scrutinization of educational inequity has been a pillar of social science research since the *Brown v. Board of Education* (1954) ruling, wherein Chief Justice Warren's opinion he stated, "segregation of white and colored children in public schools has a detrimental effect upon the colored children" and that such claims had been widely deemed credible "by [the] modern authority." Once the notion of "separate but equal" was no longer supported at the federal level the societal challenge of desegregation in public schools ensued and subsequently, the voices of "modern authority" began to increasingly feature those of social scientists.

As such, myriad studies over the years from across the social sciences have examined the extent to which educational opportunities remain restricted for disadvantaged ethnoracial groups post institutionalized integration. This dissertation argues, however, these studies often neglect to consider the dynamic ways in which social hierarchies intersect to produce masked stratification in educational contexts.

To be clear, while education in the United States has experienced formal integration, evidence of racial segregation within adolescent friendship networks in schools persists (Shrum, Cheek, and Hunter 1988; Joyner and Kao 2000). Studies that consider racial segregation in friendship networks, however, fail to capture the complex dimensionality of race in their models potentially leading to conservative estimates. These shortcomings arguably result from two insufficiencies including: 1) data limitations resulting from a lack of systematized longitudinal network data on voluntary relations in heterogenous contexts, and 2) a failure to operationalize race in its multiple dimensions, in part due to the difficulties associated with finding reliable, valid phenotypic data.

The inclination to study racial inequity in education using a unidimensional conceptualization of race is not a shortcoming limited to those investigating social networks. Indeed, most studies focused on examining how racial hierarchies structure racially disparate outcomes among students use the same approach. While studies that do consider how racial appearance impacts adolescents in schools represent an improvement in the literature, they too are limited in that they fail to consider the dynamic role race and gender play in moderating the associations between racial phenotype and educational outcomes.

To elaborate, while consensus over the persistence of racial inequality in education among researchers is for practical purposes categorical, the findings of these inquires have spurred much debate. One major point of contention emanates from disagreement over the appropriate measurement researchers are to use when operationalizing academic achievement (Kao and Thompson 2003). However, less attention has been paid to the manner in which the key explanatory variable– *racial classification*–is operationalized during model specification due to the dearth of measures aimed at capturing the multidimensionality of this complex social construct. Instead–and often resulting from the need for pragmatic research designs– researchers have overlooked the conundrum over which measures of *race* to capture during the data collection process and have opted for the measure that is most easily collected, and oftentimes, most easily analyzed. Contemporary scholarship on race, however, has argued that utilizing certain conceptualizations of the construct over others more closely aligned with the outcome under study may lead to underestimations of racial inequality, or perhaps even entirely erroneous conclusions (Bailey, Saperstein, and Penner 2014). For instance, physical phenotypic appearance has been shown to be a better predictor of outcomes related to or resulting from discrimination and prejudice. However, the ocular dimensions of race, like skin tone, are difficult to capture due to issues pertaining to their reliability and validity.

Research Themes

Therefore, in an attempt to address the abovementioned gaps my research investigates three overarching themes: 1) how best to measure differences in skin tone among individuals, 2) whether different measures of race more closely associated with others' more immediate perceptions (like measures of skin color) impact the extent of racial segregation detected in adolescent friendship networks within schools, and 3) does skin color play a role in structuring unequal educational outcomes among students in schools?

To answer these questions, I will analyze adolescent friendship network data collected at three different time points over the course of one and a half years combined with *primary* data on adolescent skin color. The data used in this analysis was derived from the Teen Identity Development and Education Study (TIDES), whose specific aims were to elucidate how ethnic-racial identity and peer relations influence the academic and social adjustment of adolescents in ethnically diverse schools. In an attempt to extend the TIDES project, a data collection design was developed to code yearbook photos of study subjects for racial phenotypic features including skin color, hair texture, nose shape, and lip shape, as well as observer racial classification. Together these measures provide a more fine-grained account of the scope of racial segregation in desegregated schools, especially considering the racialization of panethnic groups that are comprised of individuals of varying racial categories and more generally, the U.S. demographic shift toward a more multiracial population.

Why include race in its multiple dimensions?

A growing body of social scientific literature on the multidimensionality of race belabors the need for investigators to go beyond racial self-classification when conceptualizing the ways in which they are to design race-related research (Telles 2014; Jenkins 2008; Saperstein 2006). To be sure, the concept of 'race' contains distinguishable dimensions that include, how one self-identifies their own race, how others classify one's race, what one selects among limited options on a form or survey, one's phenotypic appearance as in one's skin color or other phenotypic markers, and one's racial ancestry (Roth 2016). Although the multiple dimensions of race have been shown to be significant predictors of sociological phenomenon, they are not significant uniformly. Thus, the dimension of race employed in analysis has great bearing on outcomes measuring racial stratification (Saperstein and Penner 2012; Bailey, Loveman and Muniz 2013) and should be given special theoretical consideration early on.

Social scientists interested in examining racial discrimination within academic institutions have long used racial self-classification as their key predictor. While racial self-classification is well-suited for studying demographic change and disease and illness rates, phenotypic features, such as skin color, hair texture and nose and lip shape, may be equally as important to consider when investigating racial discrimination as a social psychological phenomenon. Since the current study's aim is to investigate racially exclusionary friendship networks and racially disparate educational outcomes, and prejudice and discrimination from teachers and school administrators is one of the established factors partly driving the academic achievement gap among students (Ogbu 1986, 1992; Steele 1997; Carter 2005, Lee and Zhou 2015), measuring race in terms of skin color in addition to racial self-identification is seemingly befitting.

Road Map for Subsequent Empirical Papers

The remainder of the dissertation will be organized into three empirical papers with several theoretically grounded lines of inquiry. Paper 1 examines how to best capture skin color variation, comparing measurement across three separate skin color charts widely used throughout the social sciences. Paper 2 considers the extent of skin color homophily in adolescent friendship networks, net homophily based on racial self-identification. Paper 3 will investigate whether skin color stratification is evident across multiple outcomes of educational import and if so, whether such stratification operates differently for members of certain ethnoracial groups or genders. Finally, the dissertation will conclude with a summation of study findings and a detailed description of the direction for future research.

Introduction: References

Bailey, Stanley R., Mara Loveman, and Jeronimo O. Muniz. 2013. "Measures of 'Race' and the Analysis of Racial Inequality in Brazil." *Social Science Research* 42(1):106–19.

Bailey, Stanley, Aliya Saperstein, and Andrew Penner. 2014. "Race, Color, and Income Inequality across the Americas." *Demographic Research* 31:735–56.

Bott, Helen. 1928. "Observation of Play Activities in a Nursery School." *Genetic Psychology Monographs* 4(1):44–88.

Carter, Prudence L. 2005. *Keepin' It Real: School Success beyond Black and White*. Oxford; New York: Oxford University Press.

Hubbard, Ruth M. 1929. "A Method of Studying Spontaneous Group Formation." *Some New Techniques for Studying Social Behavior* 76–85.

Jenkins, Richard. 2014. *Social Identity*. Fourth Edition. New York: Routledge, Taylor & Francis Group.

Kao, Grace and Jennifer S. Thompson. 2003. "Racial and Ethnic Stratification in Educational Achievement and Attainment." *Annual Review of Sociology* 29(1):417–42.

Lee, Jennifer and Min Zhou. 2015. *The Asian American Achievement Paradox*. New York: Russell Sage Foundation.

Ogbu, John U. 1986. "The Consequences of the American Caste System." *The School Achievement of Minority Children: New Perspectives* 19–56.

Ogbu, John U. 1992. "Understanding Cultural Diversity and Learning." *Educational Researcher* 21(8):5–14.

Roth, Wendy D. 2016. "The Multiple Dimensions of Race." *Ethnic and Racial Studies* 39(8):1310–38.

Saperstein, Aliya. 2006. "Double-Checking the Race Box: Examining Inconsistency between Survey Measures of Observed and Self-Reported Race." *Social Forces* 85(1):57–74.

Saperstein, Aliya and Andrew M. Penner. 2012. "Racial Fluidity and Inequality in the United States." *American Journal of Sociology* 118(3):676–727.

Steele, Claude M. 1997. "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance." *American Psychologist* 52(6):613.

Telles, Edward Eric. 2014. *Pigmentocracies: Ethnicity, Race, and Color in Latin America*. 1 Edition. Chapel Hill, NC: University of North Carolina Press.

Warren, Earl. & Supreme Court of The United States. 1953. "U.S. Reports: Brown v. Board of Education." 347 U.S. 483.

Wellman, Beth. 1926. "The School Child's Choice of Companions." *The Journal of Educational Research* 14(2):126–32.

Paper 1: Examining Reliability and Validity Across Skin Color Scales and Investigating Discrimination by Color and Race

Abstract:

To disentangle some of the complexity of how racial stratification operates scholars have asserted the need for data collection efforts to expand beyond merely measuring racial self-classification but phenotypic ethnoracial markers as well. The most common way in which skin color–arguably the most salient phenotypic marker–is measured involves a procedure in which subjects' skin tone is ranked on a scale, often with a color chart as a guide. The question of how to efficiently and effectively capture skin color, however, remains due to the myriad data collection designs and color charts available to researchers aiming to capture this ocular dimension of race. Using data from the Teen Identity Development and Education Study, which coded adolescents' yearbook headshots for skin color by multiple raters, this paper asks: (1) How do three varying, widely-used skin color scales measure up in terms of their reliability and construct validity when modeling outcomes involving adolescent perceived experiences of discrimination; (2) Does one of the skin color measures provide greater unique explanatory power beyond that of racial self-identification above others when predicting different forms of discrimination; and (3) Does the predictive validity of each of the skin color scales differ for different ethnoracial groups? Results show no strong evidence that one skin color scale should be preferred over its counterparts in terms of their respective reliability and construct validity. Further, results indicate that each of the skin color scales exhibits predictive power beyond that of racial self-classification across all forms of perceived discrimination with very few statistically significant differences between color scales. Differences in predictive validity between scales among Black, Latino, Asian, and White subjects were evident. Overall scales

show the greatest predictive validity when they align best with the range of the skin tones a given ethnoracial population holds.

Introduction

The study of race and racial inequity has been a cornerstone of American sociology since its nascency (Du Bois 1903; Park et al. 1925) and the persistence and pervasiveness of racial prejudice and discrimination as a social problem in the U.S. continues to be one of the main lines of inquiry social scientists pursue (Massey and Denton 1993; Bobo 1999; Bonilla-Silva 2003). Indeed Black, Latino and Asian Americans are routinely denied access to resources, whether material or non-material, that result in racially stratified life chances that impede social mobility. However, subsumed under the process of racial discrimination is the oft overlooked problem of colorism.

Literature on colorism, defined as a discriminatory practice based on skin tone, hair texture, and facial features wherein lighter-skinned, physically Eurocentric appearing individuals are favored over their darker-skinned counterparts with fewer Eurocentric features, too has a long history albeit a less salient one. For example, ethnographers in 1941 found in small Mississippi town that light-skin and "White" type hair served as sources of prestige among Black residents, thereby making social mobility a far easier and faster process (Davis, Gardner, and Gardner 1941).

An important distinction between studies aiming to uncover the effects of colorism and those examining the effects of race and racism in the more traditional sense is that the former centers how racialized physical appearance influences unequal outcomes, while the latter does not. Thus, although the two systems of oppression often work in concert and are inextricably linked, they are indeed distinct concepts that serve to structure racial hierarchies.

While the importance of accounting for skin color in research on racial inequality has become clearer, the way of *how* to account for it remains nebulous due to the varying measurement scales utilized across studies to capture this measure. The present study aims to provide researchers with more guidance on which scale, if any, is better suited for examining the unique role skin color plays in shaping perceptions of racial discrimination among heterogenous U.S. populations. Additionally, by using an untapped source of easily accessible data–publicly available yearbooks–the present study contributes to scholarship of racial inequality among adolescents by suggesting a different method for collecting skin color date, an especially important task for those who subscribe to the multidimensional perspective of the social race construct.

Background

Why Measure Skin Tone?

The need to collect skin color data has become increasingly apparent as the number of studies evincing disparities by skin color on various outcomes of social scientific import have accumulated. Positive associations between skin color and important health outcomes like blood pressure have been replicated across several studies (Harburg et al. 1978; Laidley et al. 2019; Monk 2021). Additionally, unequal outcomes by skin color within the realms of education, income, wealth, housing, the labor market, and the criminal justice system, and perceived discrimination have all been well documented (Arce, Murguia, and Frisbie 1987; Murguia and Telles 1996; Espino and Franz, 2002; Hill 2000; Bodenhorn 2006; Gyimah-Brempong and Price 2006; Kizer 2017). Further, theorists have argued that with the increasing diversification of the not only U.S. population but global population as well, racial categorization has become more complex and, in turn, has made physical racial appearance more salient in structuring racial hierarchies (Bonilla-Silva 2004; Telles and Sue 2009). Nevertheless, the guidance on how to best capture this increasingly salient and highly meaningful measure remains limited, as various approaches continue to be employed within the field.

Approaches to Skin Color Measurement

Over the years social scientists studying skin color inequality have made several advances with regard to their data collection designs and measurement protocols for this key analytic predictor, however they may all be conceptually condensed into three overarching approaches, two of which are still widely used today in the social and psychological sciences: objective instrument evaluation, subjective evaluation using "word" categories, and subjective evaluation using color charts (Roth, 2016).

Instrument evaluation uses spectrophotometers, devices developed in the field of chemistry, to objectively measure the relative intensity of the light absorbed or reflected at a particular wavelength of light (Cary and Beckman 1941). In other words, it is a device to measure the brightness of various portions of the color spectrum. Most studies utilizing spectrophotometers measure the skin reflectance of respondents' inner arms (Monk 2015). The objectivity of color readings by spectrophotometer have been called into question when used for measuring skin tone since they are affected by foreground lighting and the part of the body targeted for measurement (Garcia and Abascal 2016).

Subjective evaluation using "word" categorical scales is a process in which survey respondents evaluate skin tone typically using a scale from 1 (very light) to 10 (very dark) without any visual guide. Categories may vary from study to study but typically follow a schema wherein lower values represent lighter pigmentation and higher values represent darker tones.

Color chart evaluation is also subjective in nature and is commonly collected through in-person interviewer classification. This in-person approach involves a matching process wherein evaluators memorize a color chart before the actual in-person interview takes place and mark the skin color they believe their respondent to be during their interaction. An increasingly common practice in studies utilizing color charts to measure skin tone provides evaluators with images of subjects and instructs raters to indicate which color in a given chart most closely resembles the skin color of the individual in the photo (Feliciano 2015).

While objective spectrophotometers are well-suited for medical studies, like dermatological studies aiming to improve understanding of how best to detect skin cancer, social classification through subjective perceptions have been argued to be more meaningful in predicting experiences of ethnoracial discrimination (Villarreal 2012). Further, skin color captured by use of categorical "word" scales, although subjective, arguably introduces so much subjectivity that the measure potentially becomes unreliable. To be sure, researchers cannot ensure that evaluators share the same notion of what constitutes light skin versus medium light skin or dark skin versus very dark skin. In fact, prior research has shown that the perceptions of skin tone may be partially dependent upon the race of the observer (Hill 2002).

Both objective color measurement by use of spectrophotometer and subjective categorical measurement without the use of color charts have been therefore viewed as less desirable methods of skin color measurement for studies on ethnoracial discrimination for seemingly diverging, yet interconnected reasons (Klonoff and Landrine 2000). While categorical classification using "word" scales may results in such unreliable measurement that the measure becomes vague, color measurement by way of an objective instrument may wash away many of the subjective underpinnings of relevance for social outcomes; thereby, circumventing many of the social under currents that make the ocular or phenotypic dimension of race a source of meaningful difference in the social world.

Studied Skin Color Scales and Potential Limitations

There are three primary skin color palettes or charts that are most ubiquitous in the social and psychological sciences–the New Immigrant Survey (NIS) skin color scale (Massey & Martin 2003), the Project on Ethnicity and Race in Latin America (PERLA) color palette (Telles 2014), and the Skin Color Assessment Procedure (SCAP) scale (adapted by Gonzales-Backen & Umaña-Taylor 2009).

The NIS skin color scale, also known as the Massey-Martin scale, is used in large-scale, nationally representative surveys such as the General Social Survey, the National Longitudinal Survey of Youth, the American national Election Studies Survey and the Fragile Families and Child Wellbeing Study. The metric is an 11-item scale ranging from zero, representing albinism or the absence of color and 10 representing the darkest possible skin. As seen in Figure 1, the scale is associated with a pictorial guide wherein each skin color is depicted on an identical male hand and shirt sleeve corresponding to the points 1 to 10 and point 0 denoting albinism is omitted. As part of the Massey-Martin procedure, interviewers were instructed to not directly compare the color chart with respondents' skin, but rather code skin color post-interview due to feelings of objectification expressed by respondents during in-person matching assessment piloting. Interviewer-reported measures of immigrant skin color in the NIS reasonably approximated skin color measures taken by spectrophotometers in the immigrants' countries of origin, thereby suggesting the scale's validity (Hersch 2008). Limitations of the NIS scale include little representation of undertones of redness and yellowness with visibly greater variation on the darker end of the color continuum, thereby potentially constraining the scale's ability to capture meaningful variation for Latino and Asian populations. Additionally, the scale has been critiqued for displaying colors depicted on male hands with shirt sleeves, as the shape of the hand and the clothing shown in the scale indicate gender and socio-economic status.

The PERLA color palette was developed as part of a two-year survey design effort for the Project on Ethnicity and Race in Latin America with help from scholars from across four countries in Latin America containing 62% of the Latin American population–Mexico, Colombia, Peru and Brazil (Telles 2014). The palette has been used in the AmericasBarometer survey, carried out by the Latin American Public Opinion Project since 2010. Similar to the NIS scale, the PERLA color palette is an 11-item chart intended to capture variations in skin color found in Latin America with an emphasis on the darker end of the color spectrum. Unlike the NIS color chart, the PERLA color palette is akin to paint samples, where each color category is represented by a rectangular color sample as seen in Figure 2. In this case, interviewers were also trained to not categorize respondents' skin tone during the interview process to avoid subjects seeing the palette, thereby reducing the likelihood of feelings of respondent objectification with the ultimate goal of increasing rapport between subject and researcher. The colors of the palette came from internet photographs complementing the self-reported measures of race and ethnicity in Latin America. The selected colors were extensively pre-tested to assess ease of use by interviewers and to examine whether they covered the range of colors found in the field.

Although the PERLA scale was developed and tested on a heterogenous population, the population was sampled in Latin America and the interviewers were also from that region of the world. While some scholars have argued that race is global in scope (Wade 2012), the region's unique historical racial projects as conceptualized by Omi and Winant (2015) influence perceptions of race and its association with racial phenotypic markers. Consequently, the palette may not be best for capturing this dimension of race among a U.S. heterogenous population with

U.S. interviewers or observers, making it potentially limiting for researchers interested in examining skin color inequity in that particular region.

The root of the third skin color measurement chart assessed in our analysis, which we call the SCAP scale, can be traced to the early 1970s (see Gitter, Mostofsky, and Satow 1972) and is arguably the procedure most widely used within the psychological sciences for skin tone measurement, although the scale itself varies widely from study to study. The procedure was later adapted by Bond and Cash (1992) resulting in a scale where colors range from 1 (very light, cream colored) to 9 (very dark, ebony). In addition, previous pilot research conducted with African-American subjects established the highly reliable, light-to-dark ordinality of the tones for the adapted procedure (Coard, Breland, and Raskin 2001). To be sure, the procedure does not include a standardized scale, but rather, investigators adapt the colors depicted on the scale and typically provide very limited information on the range of colors chosen from study to study and how the colors of a particular variation of the scale are validated. The SCAP scale used in this analysis can be seen in Figure 3 and has been adapted for a U.S. Latino/Hispanic population by Gonzales-Backen and Umaña-Taylor (2011).

The various scales utilized to measure skin color, whether standardized and used widely across studies, as is the case with the NIS and PERLA scales, or whether adapted to fit a population for one given study for one set of subjects at one given time, as is the case with the SCAP scale, beg the question of which is most capable of capturing this salient, subjectively perceived racial phenotypic feature in a meaningful way. Further, the irrefutable differences in the tones depicted in each scale (see Figures 1, 2 and 3) leave investigators interested in using a color chart for their own skin color research with the conundrum of which scale to use. It is partially this research's goal to shed light on this methodological quandary with the ultimate objective of providing more standardized guidance for those dedicated to the investigation of racial inequity by way of racial phenotypic appearance.

To our knowledge only one other study has compared the NIS and PERLA scales analytically (Gordon et al. 2022), ultimately finding sufficient consistency, comparability, and meaningfulness across instruments. While marking in important advance, the study is however limited by the number of targets rated (46 stock images) and in its ability to assess differences in predictive validity, conceptualized as meaningfulness, in that the outcome employed to test for predictive validity (i.e., perceived social experiences) are assessed by the rater. To be sure, investigators asked raters to assess the chances that the person appearing in the images to be evaluated for skin color would experience discrimination across social settings, such as in interactions with police, with healthcare, or when simply walking on the street. An arguable issue with testing the predictive validity utilizing the subjective perceptions of raters as the outcome of interest, as is the case in this study, is that it assumes evaluators are cognizant of the way skin color bias operates and believe it to be evident in the social world.

Data & Methods

To examine how to accurately capture skin color variation, this study tests the reliability and construct validity of each of the three commonly used skin color charts described above. Both reliability and validity of measurement are cornerstones to any social science research founded in empiricism. Although inextricably linked, deeming a measure reliable and deeming it valid is not synonymous. Indeed, a measure may be reliable in that it is stable and consistent, but it may not be valid in that it may lack accuracy in measuring the true construct under investigation. The distinctions between the two therefore lie in their unique definitions, which in turn demand different modes of assessment (Jones & Thissen 2006). Thus, while reliability functions as a means to make our research reproducible, validity assures that we as researchers are measuring what we intended to measure. Together, highly reliable and valid measurement produce more generalizable findings that explain meaningful variability across segments of the population.

To test for which of the three scales–the NIS skin color scale, the PERLA skin color palette, and the SCAP scale–does a more adequate job at quantifying meaningful variance in skin colors, this investigation uses primary data from the Teen Identity Development and Education Study (TIDES). TIDES involved survey data collection from the entire student body of two ethnically diverse high schools at three periods of time over a one-year period (N_{W1} = 3,191, N_{W2} = 3,605, N_{W3} = (3,109) in two distinct geographical regions of the U.S. (i.e., Southwest, Midwest). While each of the two high schools surveyed are ethnically diverse, they differ in their heterogenous composition. In each survey, participants reported on a slew of topics, including demographic characteristics, friendship networks, experiences of ethnic/racial discrimination, sense of school belonging, and depressive symptoms, among others.

Skin color data was not collected via self-report; instead, an innovative data collection design leveraging advances in computer software was developed to minimize measurement error. To be sure, while self-perceived skin color is one of the most cost-effective ways of measuring skin tone, it is not ideal for reasons formerly discussed. Additionally, interviewer classification struggles to reconcile the fact that perceptions are skewed by the individual doing the perceiving. For instance, Hill found greater variance in color measurement among Black and White interviewers when classifying respondents within their own race (2002), making the advantages of repeated measurement by multiple observers evident.

Instead, this study used subjects' color yearbook headshots for classification by multiple observers (i.e., coders). This is similar to a strategy used by other scholars (Feliciano 2015; Feliciano & Robnett 2014) using dating app profile photos. Coding of subject headshots occurred in four academic quarters–one quarter of coding for each of four yearbooks. Two quarters were completed in the winter and spring of 2019. The coding team was initially comprised of 6 female undergraduate research assistants from ethnoracially diverse backgrounds including one who identifies as "white/Caucasian," one who identifies as "Asian," two who identify as "Latino/Hispanic," one who identifies as "Latino/Hispanic" and "American Indian," and one who identifies as "Armenian." Coders were recruited from a sociology honors seminar at the University of California, Irvine. All coders received course credit for their work. The third and fourth quarters of skin color coding took place during the winter and spring of 2020 through an undergraduate research practicum. The third and fourth quarters of coding were completed by separate individuals, eight of which identified as Asian, twelve of which identified as Latino/Hispanic, three of which identified as white/Caucasian, one of which identified as Black/African American, one of which identified as Middle Eastern, and another who identified as multiracial.

To assess reliability among observers, the present study uses comparative assessment of interclass correlation coefficients across skin color chart data to denote the level of agreement between raters.

Skin Color Coding Procedure

Several steps were taken to control for factors that have been shown to influence observers' perceptions of skin tone. Matching an individual's skin color from a physical copy of a yearbook headshot to a physical color chart introduced a host of potential problems. Lighting, being one of the most pronounced factors influencing perception of color, had to be considered carefully. Coders could not code in just any environment of their choosing because this would eliminate the possibility of controlling for ambient lighting. Rather, we determined that all image coding must take place in a controlled research lab. As such, the UCI Undergraduate TIDES research lab was formed.

Additionally, extensive piloting showed that subject headshots could not appear in grid form, as in traditional yearbooks, nor could they appear in row format without impacting coder ratings. Statistical analyses of pilot data revealed that raters coded images in statistically different ways when images appeared side-by-side with other images. Thus, every yearbook page was scanned into digital form and each yearbook headshot was painstakingly cropped individually, amounting to a total of 7,735 images.

Finally, to further control for environmental factors, the subjects' skin color from their headshots had to be matched to the digital versions of the color charts under investigation, and both the subject headshots and the color charts had to appear on the same computer monitor to reduce variation resulting from differences in screen brightness and resolution. And because this particular coding task involved what can be considered objective matching, as opposed to subjective perceptions of others' race, coders were granted agency in the amount of time they were permitted to classify a target's skin color.

Skin Color as a Predictor of Perceived Discrimination to Assess Construct Validity

To illustrate the importance of accurate skin color measurement, the current study will investigate the association between skin color and perceptions of different forms of discrimination. These outcomes were intently chosen because, while different, prior research has established their association with meaningful outcomes that potentially impinge on individual life outcomes. All discrimination scales used in these analyses were developed by Fisher, Wallace and Fenton (2000).

First, overall discrimination is a measure that takes multiple forms of discrimination into account and therefore may be viewed as an indicator of how much an individual perceives that they have been discriminated against, on average. Overall perceptions of discrimination are consequential in that they have been associated with lower levels of self-esteem, higher levels of depression, and lower levels of self-efficacy (McKenzie 2006).

Peer- and school-based discrimination has long been associated with academic achievement and school discipline. For adolescents, perceived discrimination in educational settings can have important implications. For example, research indicates that discrimination within academic environments may be associated also with lower self-esteem and greater depressive symptoms, academic motivation, higher racial mistrust, and deviant behaviors (Albertini 2004; Bowman and Howard 1985; DuBois et al. 2002; Wong, Eccles, and Sameroff 2003). Research also suggests that perceptions of education-based discrimination could ultimately lead to disidentification with academics (Ogbu 1990, Steele 1997). For example, in their study of Mexican immigrant children, researchers found that peer-based discrimination was associated with more negative academic attitudes for children at moderately diverse schools (Brown and Chu 2012).

Institutional discrimination captures discrimination experienced at the hand of a store clerk, security guard, police officer, individuals working in the food service industry. Institutional discrimination has the ability to have lasting impacts on adolescent lives as it typically occurs at the hands of strangers, some of which have great authority to criminalize youth.

The present analysis builds linear regression models to assess the validity of each of the skin color measures in predicting perceptions of different forms of perceived discrimination among adolescent participants. Each model controls for survey wave, yearbook from which the subject headshot was coded for skin color, gender, grade level, highest level of parent education, and immigrant generation. These controls were chosen because they represent attributes/measures the literature suggests has bearing on the outcomes of interest. Models are stratified by region to isolate regional effects. All models were also specified to produce standardized coefficients for all continuous variables. To assess model fit, adjusted R-squares are examined.

Comparing Predictive Validity Above and Beyond Racial Self-Classification by Skin Color Scale

Subsumed under the process of racial discrimination is the oft overlooked problem of colorism (Hunter 2007). The idea that colorism operates in tandem with racism in its traditional form–that is, discrimination based on the racial classification–but cannot be subsumed by it leads one to expect valid skin color measures to have predictive power above that shown from racial self-classification measures alone. To this end, we first specify a set of models with racial self-classification alone as the key predictor. We then add each skin color measure to our models separately, and since we are comparing nested models, F-tests are run to statistically assess whether inclusion of the skin color terms show model improvement. Finally, R-squared values are compared across models. All models include the same set of controls listed previously and are also stratified by region.

Assessing Differences in Predictive Validity Between Ethnoracial Groups Across Skin Color Scales

Because the skin color scales assessed were designed for different populations, the present study examines whether evidence of varying levels of predictive validity exist across ethnoracial groups. To assess whether certain skin color scales contain greater predictive validity for certain ethnoracial groups, a set of ethnoracially stratified models including our same set of controls are specified. Standardized coefficients are compared between skin color scales within ethnoracial groups and also within skin color scales across ethnoracial groups.

Results

Reliability and Construct Validity Across Skin Color Scales

To assess interrater reliability across the NIS Skin Color Scale, the PERLA Color Palette, and the SCAP Skin Color Scale, interclass correlation coefficients (ICCs) were calculated. ICCs measure the proportion of variance that is ascribable to targets or objects of measurement (McGraw and Wong 1996). While commonly used, ICCs are of varying form and the selection of the appropriate form for interpretation is entirely dependent upon the data collection procedure and the manner in which researchers intend to apply the objects being measured. Researchers must first determine whether a) each target was rated by a different set of randomly chosen raters, b) a random sample of k raters is selected from a larger population of raters and each rater measures each target, or c) each target is rated by each of the same k raters, who are the only judges of interest (Shrout and Fleiss 1979).

While each of the three aforementioned cases require analysis of variance models to procure ICCs, their corresponding models must be specified differently. Experimental designs in which targets are rated by randomly chosen sets of raters require the implementation of one-way random effects models, where the raters are considered the random effects and results are generalizable to the larger population. Designs wherein raters are chosen from a larger population of raters with similar characteristics and each rater measures each target requires a two-way random effects model, and results are generalizable to the larger population of raters only. Lastly, in cases where targets are rated by each of the same k raters researchers must specify a two-way mixed effects model, where variance within raters are considered the random effects and variance between raters is considered the fixed effect, since they are the only raters of interest, and results are only generalizable to the fixed set of k raters.

Second, investigators must decide whether they intend to apply the measurements/ratings collected by different raters in future analysis by choosing a single rating at random or by taking the average of k ratings by each k rater (Koo and Li 2016). The present study aims to assess the interrater reliability of each of the three scales to take each to task against one another and ultimately assist potential utilizers of these scales in determining which is best suited for their investigation. To this end, we report ICCs for both the single rater case and for the average of k rater case to provide a more rigorous assessment.

The present study utilizes several coders or raters chosen at random and different sets of randomly chosen raters measured each target or subject using each of the three color scales under investigation. We therefore calculate ICCs using a one-way random effects analysis of variance model, reporting results in Table 1. Overall, each of the three skin color scales–NIS, PERLA, and SCAP–show high levels of reliability across all four yearbooks. Specifically, the interclass correlation coefficients of *k* raters shown in Table 1 reveal extremely high reliability on average for the NIS (0.98), PERLA (0.97), and SCAP (0.98) scales. Further, if we were to assume that each subject headshot was coded by a single rater, which although not the case provides a more conservative estimate of agreement, high levels of reliability remain, with the NIS producing an ICC of 0.85 on average, the PERLA producing an ICC of 0.80 on average, and the SCAP producing an ICC of 0.83 on average. Overall results reveal a nominal advantage of the NIS scale, relative to the PERLA and SCAP scales, in terms of their respective reliabilities.

(TABLE 1 ABOUT HERE)

The construct validity of the three different color scales is assessed by comparing differences in statistical significance of standardized skin color coefficients and differences in effect size for models predicting different forms of discrimination by region. In the Southwest sample, results shown in Table 2 display consistently lower standardized coefficients for the PERLA skin color measure and a non-significant coefficient when predicting peer-based discrimination. Conversely, results pertaining to the restricted Midwest sample show consistently higher effect sizes for PERLA measure. Further, the PERLA measure produces the only skin color effect that reaches statistical significance at the 0.05-level when predicting peer-based discrimination. On average, skin color measures show high construct validity when assessing different forms of discrimination. Results pertaining to associations between the PERLA scale and perceived discrimination outcomes were most distinct, but still not very different from the other two scales.

(TABLE 2 ABOUT HERE)

Differences in Increases to Explained Variance by Skin Color Measure

To assess whether one skin color scale provides greater explanatory power above racial self-classification than the others, R-squared values from models predicting overall, peer-based, school-based, and institutional discrimination are evaluated relative to R-squared values of models predicting race alone. All model control for survey wave, yearbook from which subject headshot was coded for skin color, race, gender, grade level, highest level of parent education, and immigrant generation. R-squared values were chosen because they assess the variance explained by a given model, therefore allowing one to interrogate whether the inclusion of one skin color scale over another shows a greater increase in the amount of variance explained. Additionally, F-tests reveal whether the inclusion of each of the skin color measures produces better fitting models than those including race alone. Results shown in Table 3 display results by school region to isolate potential regional effects.

The three measures of skin color differ only minimally in the additional variance they explain over and above race alone. Models restricted to the Southwest analytic sample reveal no differences in the increase in variance explained when predicting overall discrimination (R^2's =0.12 for NIS, PERLA and SCAP skin color measures) and institutional-based discrimination (R^2's =0.20 for NIS, PERLA and SCAP skin color measures). Models predicting peer-based discrimination among the Southwest sample show no increase in variance explained when adding the PERLA skin color measure to the race only model, as R-squared value for the race only model is equal to that produced from the model that also includes the PERLA skin color measure (R^2's =.04); the corresponding F-statistic of 1.8 is also not significant ($p>.05$). Models predicting school-based discrimination among the Southwest sample show an increase in variance explained when adding the NIS, PERLA, and SCAP skin color measure to the race only model as the F-statistics testing the difference between the race only model and the models including each of the skin color measures are all significant at least the 0.05-level. While an increase in variance explained is evident at the separate inclusion of each of the skin color measures, the greatest increase stems from the model including the NIS skin color measure. In no case, however, is the increase in the R-squared very large when adding any of the skin color measures.

Results pertaining to the Midwest sample vary somewhat from those pertaining to the Southwest sample. Like models restricted to the Southwest sample, models restricted to the Midwest analytic sample reveal no differences in the increase in variance explained when predicting institutional-based discrimination (R^2's =0.23 for NIS, PERLA and SCAP skin color measures). However, unlike results pertaining to the Southwest sample, models predicting peer-based discrimination among the Midwest sample show the only statistically significant increase in variance explained exist in the model that includes the PERLA skin color measure. To be sure, while each of the R-squared values equal 0.07, the F-statistic of 4.9 is the only significant statistic at the 0.05-level. Models predicting overall discrimination also show a slight advantage for the inclusion of the PERLA measure (R^2's =0.16) over the NIS (R^2's =0.15) and SCAP (R^2's =0.15) measures, while models predicting school-based discrimination show a smaller increase in variance explained stemming from inclusion of the SCAP measure (R^2's =0.09), relative to the increase stemming from the inclusion of the NIS (R^2's =0.10) and PERLA (R^2's =0.10) measures.

Overall, results presented in Table 3 reveal that for the Southwest region there appears to be a slight advantage to the inclusion of the NIS skin color measure, on average, while within the Midwest region a slight on average advantage is apparent for the inclusion of the PERLA skin color scale. Differences may stem from differences in composition of the ethnoracial diversity between regions.

(TABLE 3 ABOUT HERE)

Difference in Predictive Validity by Ethnoracial Group

Table 4 presents estimates of associations between subject skin tone and overall, peer-based, school-based, and institutional perceived discrimination for self-identified Black, Latino, Asian, and White subsamples.

When examining differences in effect sizes within ethnoracial groups, it is evident that among the Black subsample the SCAP skin color measure has consistently lower predictive power across models predicting overall (ß =0.135), peer-based (ß =0.090), school-based (ß =0.090), and institutional (ß =0.0.162) discrimination due to consistently lower effect sizes relative to those attributed to the NIS and PERLA skin color measures.

Results pertaining to the Latino subsample show statistical significance for the PERLA skin color measure when estimating associations with overall (ß =0.082, $p<0.05$) and institutional (ß =0.106, $p<0.05$) discrimination and also for the NIS measure when predicting institutional discrimination (ß =0.086, $p<0.05$). Inclusion of the SCAP skin color measure across all models reveal no significant associations between skin color and discrimination for the Latino subsample.

Models restricted to the Asian subsample only reveal statistical significance for the NIS skin color measure when predicting school-based discrimination (ß =0.105, $p<0.05$). That no other skin color measure was statistically significantly associated with any of the measures of discrimination? suggests that skin color, whether measured using any of the three scales under investigation, is less consequential in determining experiences of discrimination among Asians.

When examining models estimating different forms of discrimination by skin color measure using the White subsample, there is evidence of higher predictive validity when utilizing the SCAP skin color measure, relative to the NIS and PERLA measures. To be sure, while each of the skin color measures produce statistically significant effects when predicting overall and institutional discrimination at the 0.05-level, and no statistical significance when predicting peer-based discrimination, results show differences in statistical significance when predicting school-based discrimination, with only the NIS ($ß_{NIS}$=0.062, $p<0.05$) and SCAP ($ß_{SCAP}$=0.063, $p<0.05$) measures reaching significance. That the SCAP measure produces higher effect sizes throughout models predicting all for types of discrimination suggests higher predictive validity, relative to the other skin color measures, for the White subjects.

In sum, results evince predictive validity across all three skin color scales when examining perceptions of discrimination among Black subjects, however, effect sizes were consistently smaller for SCAP scale measures. Predictive validity amongst Latino subjects is highest when utilizing the PERLA scale to capture skin color as suggested by the measure's greater number of significant coefficients coupled with larger effect sizes. The NIS skin color measure shows greater predictive validity among Asian subjects, as it was the only skin color measure that produced a statistically significant coefficient, albeit only for models predicting school-based discrimination. Finally, the SCAP measure shows higher predictive validity among White subjects, as seen in consistently significant results throughout models predicting different forms of discrimination (even though only reaching marginal significance when assessing peer-based discrimination) and in constantly larger effect sizes across all discrimination outcomes.

(TABLE 4 ABOUT HERE)

Discussion

Overall, high reliability and similar construct validity show that despite their marked differences at face value (see Figures 1 and 2), any of the scales can be used confidently by researchers interested in investigating the effects of skin color inequality among diverse populations. Specifically, ICC results demonstrated that, in a controlled environment, different raters provided consistent ratings of the same subject headshot with the NIS, PERLA and SCAP scales. While most scales displayed construct validity in both regions, the exception of the PERLA scale in the Southwest region and the exception of the NIS and SCAP scales in the Midwest region suggest the importance of considering contextual factors when deciphering which scale to use in the data collection process, as choosing one scale over another has the potential to yield varying results. The importance for consideration of contextual factors is further supported by results showing differences to increases in explained variance relative to models including self-identified race plus controls alone. To be sure, that the inclusion of the NIS scale produced the largest increases to explained variance in the Southwest region, on average, and that the inclusion of the PERLA scale produced the largest increases in the Midwest region shows that the greater variation in darker skin tones inherent in the NIS scale may be better at explaining the effect of colorism in regions where black individuals represent a smaller proportion of the population relative to other minorities of color, while the greater variation in lighter skin tones inherent in the PERLA scale may be better at explaining the same effect in regions where minorities with lighter skin tones, such as the Latino population in the Midwest, represent a smaller proportion of the minority population.

Ethnoracially stratified models evaluating predictive validity by skin color scale show predictive power amongst all of the skin color scales for the Black analytic sample, with consistently lower effect sizes when the SCAP scale is used. Findings are likely due to visible limited variation on the darker end of the spectrum for this given scale, thereby erasing meaningful variance among Black subjects. That models amongst the Latino subsample show more statistical significance, on average, when utilizing the PERLA scale are likely due to greater variation on the lighter and medium end of the spectrum capturing meaningful variation for this particular subgroup that tends to have a greater proportion of its population with lighter and medium skin tones. Overall scales show the greatest predictive validity when they align best with the range of the skin tones a given ethnoracial population holds. The SCAP showed greater predictive validity among White subjects, again likely due to the greater variation inherent in the scale's lighter end of the spectrum. Finally, the Asian subsample revealed no predictive validity when assessing different forms of discrimination, aside for those predicting school-based discrimination and only when using the NIS scale. Researchers should therefore strongly consider the population under study when deciding which skin color scale to use in data collection.

Skin color is significantly associated with every form of discrimination and this association cannot be explained away by self-identified race. Findings are in line with research on the multidimensionality of race and add further evidence in support of race scholars' calls to consider differing aspects of race, particularly those of phenotypic appearance, in investigations of racial inequity (Roth 2016).

Lastly, results show the association between skin color and perceived discrimination is weaker on average for Asian students, regardless of the scale used to measure skin color. Findings align substantively with prior research on stereotype promise, wherein Asian-Americans, many of darker skin tones and of South-East Asian descent, procure advantage from being associated with high-achieving, light-skinned Asians of East-Asian descent (Lee and Zhou 2015).

The strengths of this study include the use of a large phenotypic dataset containing three separate skin color measures collected with the use of two of the most widely used skin color scales as guides and one scale adapted for a lighter population (i.e., SCAP), and the availability of multiple measures of perceived discrimination combined with large enough subsamples of Black, Latino, Asian, White subjects that allowed for comparison of scales between ethnoracial groups.

Paper 1: References

Albertini, Velmarie L. 2004. “Racial Mistrust among Immigrant Minority Students.” *Child and Adolescent Social Work Journal* 21 (4): 311–31.

Arce, Carlos H., Edward Murguia, and W. Parker Frisbie. 1987. “Phenotype and Life Chances Among Chicanos.” *Hispanic Journal of Behavioral Sciences* 9 (1): 19–32.

Bailey, Stanley R., Fabrício M. Fialho, and Andrew M. Penner. 2016. “Interrogating Race: Color, Racial Categories, and Class Across the Americas.” *American Behavioral Scientist* 60(4):538–55.

Bailey, Stanley R, Mara Loveman, and Jeronimo O Muniz. 2013. “Measures of ‘Race’ and the Analysis of Racial Inequality in Brazil.” *Social Science Research* 42 (1): 106–19.

Bailey, Stanley, Aliya Saperstein, and Andrew Penner. 2014. “Race, Color, and Income Inequality across the Americas.” *Demographic Research* 31: 735–56.

Banton, Michael. 2012. “The Colour Line and the Colour Scale in the Twentieth Century.” *Ethnic and Racial Studies* 35(7):1109–31.

Bobo, Lawrence D. 1999. “Prejudice as Group Position: Microfoundations of a Sociological Approach to Racism and Race Relations.” *Journal of Social Issues* 55 (3): 445–72.

Bodenhorn, Howard. 2006. “Colorism, Complexion Homogamy, and Household Wealth: Some Historical Evidence.” *American Economic Review* 96 (2): 256–60.

Bond, Selena, and Thomas F. Cash. 1992. “Black Beauty: Skin Color and Body Images among African-American College Women.” *Journal of Applied Social Psychology* 22 (11): 874–88.

Bonilla-Silva, Eduardo. 2003. *Racism without Racists: Color-Blind Racism and the Persistence of Racial Inequality in the United States*. Lanham, Md: Rowman & Littlefield.

Bowman, Phillip J, and Cleopatra Howard. 1985. “Race-Related Socialization, Motivation, and Academic Achievement: A Study of Black Youths in Three-Generation Families.” *Journal of the American Academy of Child Psychiatry* 24 (2): 134–41.

Brown, Christia Spears and Hui Chu. 2012. “Discrimination, Ethnic Identity, and Academic Outcomes of Mexican Immigrant Children: The Importance of School Context.” *Child Development* 83(5):1477–85.

Cary, HH and Arnold O. Beckman. 1941. “A Quartz Photoelectric Spectrophotometer.” *JOSA* 31(11):682–89.

Coard, Stephanie Irby, Alfiee M. Breland, and Patricia Raskin. 2001. “Perceptions of and Preferences for Skin Color, Black Racial Identity, and Self-Esteem Among African Americans1.” *Journal of Applied Social Psychology* 31(11):2256–74.

Davis, A, BB Gardner, and MR Gardner. 1941. “Deep South; a Social Anthropological Study of Caste and Class.”

Du Bois, W.E.B. (1903) 1999. *The Souls of Black Folk*, edited by Henry Louis Gates, Jr. and Terri Hume Oliver. New York: W.W. Norton.

DuBois, David L, Carol Burk–Braxton, Lance P Swenson, Heather D Tevendale, and Jennifer L Hardesty. 2002. “Race and Gender Influences on Adjustment in Early Adolescence: Investigation of an Integrative Model.” *Child Development* 73 (5): 1573–92.

Espino, Rodolfo, and Michael M Franz. 2002. “Latino Phenotypic Discrimination Revisited: The Impact of Skin Color on Occupational Status.” *Social Science Quarterly* 83 (2): 612–23.

Feliciano, Cynthia. 2015. “Shades of Race: How Phenotype and Observer Characteristics Shape Racial Classification.” *American Behavioral Scientist* 60(4):390–419.

Feliciano, Cynthia and Belinda Robnett. 2014. “How External Racial Classifications Shape Latino Dating Choices.” *Du Bois Review: Social Science Research on Race* 11(2):295–328.

Fisher, Celia B, Scyatta A Wallace, and Rose E Fenton. 2000. “Discrimination Distress during Adolescence.” *Journal of Youth and Adolescence* 29 (6): 679–95.

Garcia, Denia and Maria Abascal. 2015. “Colored Perceptions: Racially Distinctive Names and Assessments of Skin Color.” *American Behavioral Scientist* 60(4):420–41.

Gitter, A. George, David I. Mostofsky, and Yoichi Satow. 1972. “The Effect of Skin Color and Physiognomy on Racial Misidentification.” *The Journal of Social Psychology* 88(1):139–43.

Gonzales-Backen, Melinda A. and Adriana J. Umana-Taylor. 2011. “Examining the Role of Physical Appearance in Latino Adolescents’ Ethnic Identity.” *Journal of Adolescence* 34(1):151–62.

Gordon, Rachel A., Amelia R. Branigan, Mariya Adnan Khan, and Johanna G. Nunez. 2022. “Measuring Skin Color: Consistency, Comparability, and Meaningfulness of Rating Scale Scores and Handheld Device Readings.” *Journal of Survey Statistics and Methodology* 10 (2): 337-64.

Gyimah-Brempong, Kwabena, and Gregory N. Price. 2006. “Crime and Punishment: And Skin Hue Too?” *American Economic Review* 96 (2): 246–50.

Harburg, Ernest, Lillian Gleibermann, Peter Roeper, M Anthony Schork, and William J Schull. 1978. “Skin Color, Ethnicity, and Blood Pressure I: Detroit Blacks.” *American Journal of Public Health* 68 (12): 1177–83.

Hersch, Joni. 2008. “Skin Color, Immigrant Wages, and Discrimination.” Pp. 77–90 in *Racism in the 21st Century*. Springer.

Hill, Mark E. 2000. "Color Differences in the Socioeconomic Status of African American Men: Results of a Longitudinal Study." *Social Forces* 78 (4): 1437–60.

Hill, Mark E. 2002. "Race of the Interviewer and Perception of Skin Color: Evidence from the Multi-City Study of Urban Inequality." *American Sociological Review* 67(1):99–108.

Hunter, Margaret. 2007. "The Persistent Problem of Colorism: Skin Tone, Status, and Inequality." *Sociology Compass* 1 (1): 237–54.

Jones, Lyle V. and David Thissen. 2006. "1 A History and Overview of Psychometrics." Pp. 1–27 in *Handbook of Statistics*. Vol. 26, edited by C. R. Rao and S. Sinharay. Elsevier.

Kelly, Monique DA. 2020. "Examining Race in Jamaica: How Racial Category and Skin Color Structure Social Inequality." *Race and Social Problems* 12 (4): 300–312.

Kizer, Jessica M. 2017. "Arrested by Skin Color: Evidence from Siblings and a Nationally Representative Sample." *Socius: Sociological Research for a Dynamic World* 3:1-12

Koo, Terry K, and Mae Y Li. 2016. "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research." Journal of Chiropractic Medicine 15 (2): 155–63.

Laidley, Thomas, Benjamin Domingue, Piyapat Sinsub, Kathleen Mullan Harris, and Dalton Conley. 2019. "New Evidence of Skin Color Bias and Health Outcomes Using Sibling Difference Models: A Research Note." *Demography* 56 (2): 753–62.

Lee, Jennifer, and Min Zhou. 2015. *The Asian American Achievement Paradox*. New York: Russell Sage Foundation.

Massey, Douglas S, and Nancy A Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press.

Massey, Douglas S., and Jennifer A. Martin. 2003. The NIS Skin Color Scale.

McGraw, Kenneth O., and S. P. Wong. 1996. “Forming Inferences about Some Intraclass Correlation Coefficients.” Psychological Methods 1 (1): 30–46.

McKenzie, Kwame. 2006. “Racial Discrimination and Mental Health.” *Psychiatry* 5 (11): 383–87.

Monk, Ellis P., Jr. 2016. “The Consequences of ‘Race and Color’ in Brazil.” *Social Problems* 63 (3): 413–30.

Monk, Ellis P. 2021. “Colorism and Physical Health: Evidence from a National Survey.” *Journal of Health and Social Behavior* 62 (1): 37–52.

Murguia, Edward, and Edward E. Telles. 1996. “Phenotype and Schooling among Mexican Americans.” *Sociology of Education*, 276–89.

Park, Robert E., Ernest W. Burgess, and Roderick D. McKenzie, and Louis Wirth. 1925. *The City*. Chicago, Ill: The University of Chicago Press.

Perreira, Krista M., and Edward E. Telles. 2014. “The Color of Health: Skin Color, Ethnoracial Classification, and Discrimination in the Health of Latin Americans.” *Social Science & Medicine* 116 (September): 241–50.

Ogbu, John U. 1990. “Literacy and Schooling in Subordinate Cultures: The Case of Black Americans.” *Going to School: The African American Experience*, 113–31.

Omi, Michael and Howard Winant. 2014. *Racial Formation in the United States*. Routledge.

Shrout, Patrick E., and Joseph L. Fleiss. 1979. “Intraclass Correlations: Uses in Assessing Rater Reliability.” Psychological Bulletin 86 (2): 420–28.

Steele, Claude M. 1997. “A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance.” *American Psychologist* 52 (6): 613.

Roth, Wendy D. 2016. “The Multiple Dimensions of Race.” *Ethnic and Racial Studies* 39 (8): 1310–38.

Telles, Edward Eric. 2014. *Pigmentocracies: Ethnicity, Race, and Color in Latin America*. 1 Edition. Chapel Hill, NC: University of North Carolina Press.

Wade, Peter. 2012. “Skin Colour and Race as Analytic Concepts.” *Ethnic and Racial Studies* 35 (7): 1169–73.

Wong, Carol A, Jacquelynne S Eccles, and Arnold Sameroff. 2003. “The Influence of Ethnic Discrimination and Ethnic Identification on African American Adolescents’ School and Socioemotional Adjustment.” *Journal of Personality* 71 (6): 1197–1232.

Table 1. Interclass correlation coefficients measuring interrater reliability across 3 skin color scales using one-way random effects analysis of variance models

	ICC (1,1)	ICC (1, k)	Total Number of Subjects	Total Number of Coders
NIS				
Southwest, Year 1	0.87***	0.98***	2,686	6
Southwest, Year 2	0.82***	0.98***	2,814	13
Midwest, Year 1	0.81***	0.98***	1,597	12
Midwest, Year 2	0.89***	0.98***	1,610	5
NIS average	0.85	0.98	–	–
PERLA				
Southwest, Year 1	0.75***	0.95***	2,686	6
Southwest, Year 2	0.84***	0.98***	2,814	13
Midwest, Year 1	0.82***	0.98***	1,597	12
Midwest, Year 2	0.80***	0.95***	1,610	5
PERLA average	0.80	0.97	–	–
SCAP				
Southwest, Year 1	–		–	–
Southwest, Year 2	0.82***	0.98***	2,814	13
Midwest, Year 1	0.82***	0.98***	1,597	12
Midwest, Year 2	0.86***	0.97***	1,610	5
SCAP Average	0.83	0.98	–	–

Source: Teen Identity Development and Education Study
Notes: * p<0.05; ** p<0.01; *** p<0.001. Each coder did not rate each target, except for the Michigan 2017-2018 yearbook. The Michigan 2017-2018 yearbook therefore uses two-way mixed effects models to calculate ICCs.

Table 2. Linear regression models predicting frequency of perceived overall, peer-based, school-based, and institutional discrimination

	Southwest			**Midwest**		
	Estimate	***p***	**(SE)**	**Estimate**	***p***	**(SE)**
Overall discrimination						
NIS skin color	0.176	***	(0.04)	0.168	***	(0.04)
PERLA skin color	0.127	***	(0.04)	0.193	***	(0.04)
SCAP skin color	0.161	***	(0.04)	0.116	**	(0.04)
Peer-based discrimination						
NIS skin color	0.101	*	(0.04)	0.075	†	(0.04)
PERLA skin color	0.050		(0.04)	0.097	*	(0.04)
SCAP skin color	0.107	**	(0.04)	0.027		(0.04)
School-based discrimination						
NIS skin color	0.110	**	(0.04)	0.197	***	(0.04)
PERLA skin color	0.079	*	(0.04)	0.221	***	(0.04)
SCAP skin color	0.087	*	(0.04)	0.155	***	(0.04)
Institutional discrimination						
NIS skin color	0.226	***	(0.04)	0.200	***	(0.04)
PERLA skin color	0.187	***	(0.03)	0.213	***	(0.04)
SCAP skin color	0.202	***	(0.04)	0.149	***	(0.04)
N	1,953		–	1,536		–

Source: Teen Identity Development and Education Study
Notes: * p<0.05; ** p<0.01; *** p<0.001. All models control for survey wave, yearbook from which subject headshot was coded for skin color, race, gender, grade, highest level of parent education, and immigrant generation. All skin color coefficients are standardized. Models may have somewhat different cases due to missingness.

Table 3. Adjusted R-Squares for models predicting overall, peer-based, school-based, and institutional discrimination by school region

	Southwest			Midwest		
	Race	Race & Skin Color	F-statistic	Race	Race & Skin Color	F-statistic
Overall discrimination	0.11			0.14		
NIS		0.12	21.1***		0.15	16.2***
PERLA		0.12	12.8***		0.16	21.3***
SCAP		0.12	18.7***		0.15	7.9**
Peer-based discrimination	0.04			0.069		
NIS		0.05	6.4*		0.07	3.0†
PERLA		0.04	1.8		0.07	4.9*
SCAP		0.05	7.6**		0.07	0.4
School-based discrimination	0.05			0.08		
NIS		0.06	7.8**		0.10	20.95***
PERLA		0.05	4.7*		0.10	26.6***
SCAP		0.05	5.2*		0.09	13.3***
Institutional discrimination	0.19			0.22		
NIS		0.20	38.0***		0.23	25.5***
PERLA		0.20	30.0***		0.23	28.3***
SCAP		0.20	31.9***		0.23	14.1***
N		1,953			1,536	

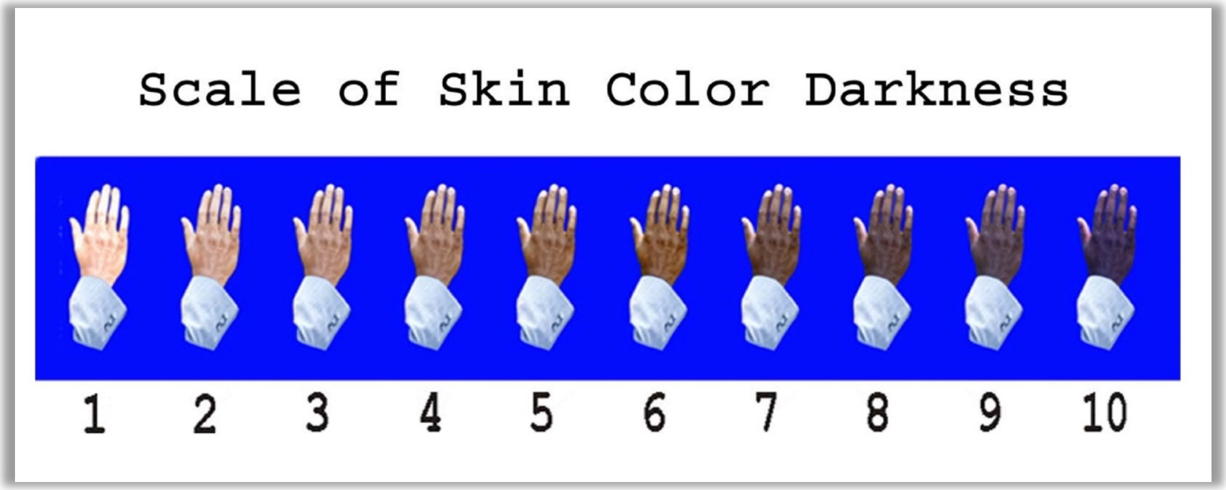
Source: Teen Identity Development and Education Study
Notes: * p<0.05; ** p<0.01; *** p<0.001. All models control for survey wave, yearbook from which subject headshot was coded for skin color, gender, grade, highest level of parent education, and immigrant generation. Models may have somewhat different cases due to missingness.

Table 4. OLS regression models estimating adolescent perceived overall, peer-based, school-based, and institutional discrimination as a function of NIS skin color, PERLA skin color, and SCAP skin color by ethnoracial group

	Black (N=839)		Latino (N=669)		Asian (N=468)		White (N=1,359)	
	ß	**SE**	**ß**	**SE**	**ß**	**SE**	**ß**	**SE**
Overall discrimination								
NIS	0.164	(0.036)***	0.077	(0.042)†	0.048	(0.050)	0.083	(0.028)**
PERLA	0.170	(0.037)***	0.082	(0.042)*	0.028	(0.051)	0.057	(0.029)*
SCAP	0.135	(0.037)***	0.080	(0.044)†	0.015	(0.051)	0.100	(0.028)***
Peer-based discrimination								
NIS	0.127	(0.037)***	0.045	(0.042)	-0.023	(0.049)	0.028	(0.028)
PERLA	0.121	(0.038)**	0.037	(0.041)	-0.039	(0.050)	0.008	(0.029)
SCAP	0.090	(0.037)*	0.062	(0.043)	-0.043	(0.050)	0.054	(0.028)†
School-based discrimination								
NIS	0.114	(0.036)**	0.076	(0.043)†	0.105	(0.050)*	0.062	(0.028)*
PERLA	0.124	(0.037)***	0.074	(0.042)†	0.095	(0.051)†	0.050	(0.029)†
SCAP	0.090	(0.036)*	0.066	(0.044)	0.069	(0.051)	0.063	(0.028)*
Institutional discrimination								
NIS	0.179	(0.036)***	0.086	(0.043)*	0.091	(0.051)†	0.140	(0.028)***
PERLA	0.192	(0.037)***	0.106	(0.042)*	0.067	(0.051)	0.106	(0.029)***
SCAP	0.162	(0.037)***	0.080	(0.044)†	0.055	(0.051)	0.147	(0.028)***

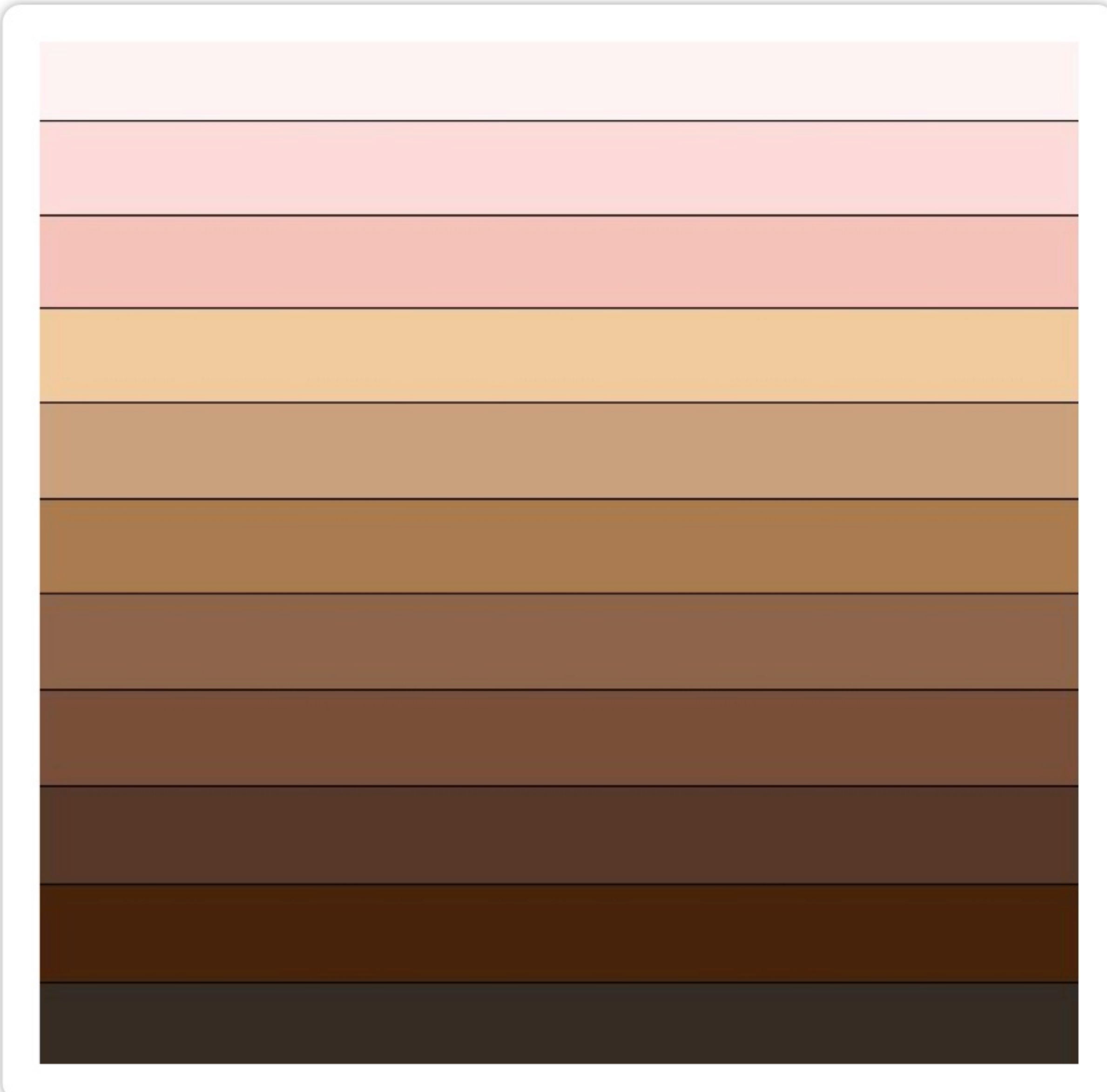
Source: Teen Identity Development and Education Study.
Notes: * p<0.05; ** p<0.01; *** p<0.001. All models control for survey wave, yearbook from which subject headshot was coded for skin color, gender, grade, highest level of parent education, and immigrant generation. Coefficients are standardized. Models may have somewhat different cases due to missingness.

Figure 1. NIS Skin Color Scale



Source: Massey, Douglas S., and Jennifer A. Martin. 2003. The NIS Skin Color Scale.

Figure 2. PERLA Skin Color Palette



Source: Telles, Edward Eric. 2014. Pigmentocracies: Ethnicity, Race, and Color in Latin America. 1 Edition. Chapel Hill, NC: University of North Carolina Press.

Figure 3. SCAP Color Guide

Source: Gonzales-Backen, Melinda A, and Adriana J Umana-Taylor. 2011. "Examining the Role of Physical Appearance in Latino Adolescents' Ethnic Identity." Journal of Adolescence 34 (1): 151–62.

Paper 2: Above and Beyond Racial Homophily: How Overlooked Skin Color Stratification Masks Racial Segregation in Adolescent Friendship Networks

Abstract:

While former studies lay a strong foundation for the investigation of racial exclusionary networks, they neglect a body of literature on colorism and phenotype-ism, which argues that skin color discrimination and discrimination based on other phenotypic features such as hair texture, nose shape and lip shape play a key role in structuring racial hierarchies. Using data from the Teen Identity Development and Education Study, the present study therefore uses Exponential Random Graph Models (ERGMs) to assess whether those similar in skin color are more likely to be tied to one another within reciprocal friendships, whether this pattern holds net of homophily based on racial self-classification, and if so, whether such patterns are more evident within some ethnoracial groups than others. Findings show evidence of skin color homophily above and beyond that of racial homophily. Significant differences in the extent of skin color homophily by ethnoracial group were also evident. The study's findings underscore the importance of operationalizing race in multiple dimensions when assessing racial exclusion in social networks to avoid yielding conservative estimates.

Introduction

While past research has demonstrated that group formation is highly predicated on race, little is known about the cleavages embedded *between* and *within* ethnoracial groups through phenotypic difference. Homophily—the idea that similarity breeds connection—has been a guiding principle in the study of social relations dating back to Ancient Greece. This principle—the homophily principle—structures network ties of every type including marriage, information exchange, organizational comembership, and friendship, among others (Blau 1977; McPherson, Smith-Lovin, and Cook 2001). More importantly, the tendency to interact with people similar to oneself is of special concern to social scientists because such biases or inclination perpetuate inequality by limiting access to resources and information among the less advantaged through a process of cumulative disadvantage (Merton 1968).

Past research on racial and ethnic homophily is argued to be the strongest type of division within networks in American society (Kao and Joyner 2004; Kalmijn 1998; Marsden 1987; 1988). Several studies have documented this tendency across several types of relations. For instance, in his study on the "discussion of important matters" (i.e., core discussion networks), Marsden found that only 8% of adults with networks size two or more mention doing so with an individual of another race—less than one seventh the diversity we would expect if people chose from the population randomly (1988). Other work centered on work and organization found that primary or intimate work ties were highly racially homogenous (Lincoln and Miller 1979; Ibarra 1995). Among adolescents in particular the research is not scant either (Shrum, Cheek, and Hunter 1988; Joyner and Kao 2000; Moody 2001; Goodreau, Kitts and Morris 2009). While these studies certainly have advanced understanding of how race structures social networks, few have conceptualized race as a multidimensional social construct irreducible to mere racial self-classification.

Indeed, an extent literature on colorism or phenotypic stratification argues that skin color discrimination and discrimination based on other phenotypic features such as hair texture, nose shape, lip shape, and eye color play a key role in structuring racial hierarchies. For instance, skin color inequity has been shown to manifest itself early on in the life course in the form of skin color preferences among school children (Hughes and Hertel 1990; Anderson and Cromwell 1977). Further, eminent race theorists predict that drastic changes in U.S. demography are increasing the importance of racial phenotype in determining life chances for U.S. societal members (Bonilla-Silva 2004).

Specifically, scholars argue that due to the rapid increase in the non-white population, a whitening of the population through immigration and a broadening of the white racial category to include those hierarchically closest will ensue as a means to maintain white power (Helg 1990). Historically this process has taken place, as is evinced by the reluctant inclusion of the Irish, Italian and American Jewish population within the white racial strata (Waters 1990; Roediger 1999; Portes and Rumbaut 2004) and more recent empirical findings suggest this process may be presently occurring, as evidenced by the increasing rates of Asian-white and Latino-white marriage unions (Moran 2003) and the lower rates of residential segregation from whites among light-skinned Latinos and Asians, relative to blacks (Massey and Denton 1987). On the other hand, dark-skinned Latinos have been shown to experience similar residential segregation rates as blacks (Bonilla-Silva 2004). Together these findings suggest that, if found, the extent of skin color homophily within adolescent friendship networks may be conditioned on the ethnoracial identification of the individuals engaged in a friendship tie.

This research therefore aims to bridge the well-known literature on skin color inequality with the social network analysis approach to examine the extent to which racial phenotype structures friendship ties among adolescents in schools. Synthesizing these ideas, the proposed study asks*: 1) Are adolescents similar in skin color more likely to be tied to one another within friendships, net of homophily based on racial self-classification and 2) Is there evidence that the extent of skin color homophily differs across ethnoracial groups?* The answers to these questions help elevate our understanding of the complex state of race relations among individuals' voluntary friendship networks, the extent to which informal segregation through voluntary friendships in schools persists, and signal whether a broadening of the white ethnic category is materializing.

The empirical chapter will be organized in the following way: first, an overview of the sociological theoretical advances motivating the present analysis will be discussed, followed by a more detailed review of the literature on skin color stratification within the educational context. Next a set of hypotheses will be presented, and a detailed description of the data and methodology to be implemented will be outlined. Finally, analytic results will be interpreted, and a discussion of their substantive significance will ensue.

Background

Symbolic Interactionism and Racial Stratification: How Racial Hierarchies Shape and are Shaped by Human Interaction

Since his initial writings on social order, the founding father of systematized sociological inquiry, Émile Durkheim, theorized the existence of a social phenomenon characterized as the degree to which members of a given group share common sentiments and beliefs–a collective consciousness (Durkheim 1893). Further, he argued that this phenomenon was what he termed a social fact, operating external to individuals and exerting constraint over them. In contrast to the emphasis Durkheim placed on cooperative interdependence and solidarity, Weber advanced the field of sociology by arguing that social life included not only cooperation among societal members but also exploitation and conflict between them (Coser 1971; Johnson 1981). Of all classical theorists, however, Simmel was the first to insist that society cannot exist independent of the process of human interaction.

By the time sociology reached its rapid expansion in the mid-twentieth century, scholars like Mead and Cooley, and students Blumer and Goffman, further refined the field with their formulation and specification of symbolic interactionism, whose underlying assumptions assert individuals construct meaning via the interaction process, the self-concept (which is influenced by others) is a motivation for behavior, and a unique, dialectic relationship exists between the individual and society. To be sure, theories within the realm of symbolic interactionism underscored the important role micro-level interaction plays in making up social structure and how social structure in turn impacts and potentially constrains human interaction (Mead 1934; Blumer 1958; Cartwright 1979). Together these contributions lay the bedrock for the investigation of relationships constrained by racialized similarity. Simply put, while racial hierarchies influence who individuals choose to interact with and befriend, who individuals befriend influence individual life chances, thereby evincing the dialectic feedback loop between social structure and human interaction that symbolic interactionists theorize.

Theories of Racial Group Position and Prejudice

A large segment of theory on prejudicial attitudes figures prominently within social psychology as practiced by those trained primarily in psychology (Katz & Braly 1933, 1935; Adorno et al. 1950; Allport 1954). It has been suggested, however, that viewing racial prejudice through a more sociological lens potentially holds advantages because it necessitates consideration of how social structure comes to shape individual psychology and socially consequential behavior, thereby allowing for a more synthetic approach (Bobo 1999). Some of the initial theories of prejudice founded within the sociological tradition may be traced back to Herbert Blumer's group position model (1958), wherein he proposed that prejudice was inextricably linked to "sense of group position" that had enduring collective properties to it. To be clear, Blumer maintained that racial attitudes or prejudice should be conceptualized as general orientation containing normative ideas of where one's own group should stand relative to outgroup members and these normative ideas, in turn, influence whom and how individuals interact with one another. Although Blumer provided a non-reductionist lens in his postulation of racial attitudes, he did not elaborate on the fluid nature of race and its multidimensional definitions and meanings.

The Ocular Nature of Race and Its Increasing Significance

Although contemporary sociologists dedicated to the study of race and ethnicity have long contended that race is a social construct rather than a biological one, few have extended their work to consider racial phenotypic appearance and its relative importance in determining socially constructed racial hierarchies. Better stated by Omi and Winant in their pivotal work on the topic, they argue:

"[t]here is a crucial *corporeal* dimension to the race-concept. Race is *ocular* in an irreducible way. Human bodies are visually read, understood, and narrated by means of symbolic meanings and associations. Phenotypic differences are not necessarily seen or understood in the same consistent manner across time and place, but they are nevertheless operating in specific social settings. Not because of any biologically based or essential difference among human beings across such phonemic variables as "color" or "hair texture," but because such sociohistorical practices as conquest and enslavement classified human bodies for purposes of domination as well–racial phenotypes such as black and white have been constructed and encoded through the language of race. We define this process as *racialization [...] (2015:13)"*

In addition, some work suggests that the former black-white divide characterizing the U.S. racial hierarchy is shifting towards a more complex racial order that is more concerned with physical racialized presentation due to demographic changes; thereby predicting its increasing importance in social life (Bonilla-Silva 2002). For instance, scholars have shown how skin color influences various outcomes among individuals including those centered on education (Allen, Telles, and Hunter 2000), the labor market (Espino and Franz 2002; Kreisman and Rangel 2015), housing (Yinger 1991; Hakken 1979; Denton and Massey 1989), spousal status and dating patterns (Edward, Carter-Tellison, Herring 2004; Feliciano and Robnett 2014), criminal justice (Blair, Judd and Chapleau 2004; Burch 2005), and mental health (Codina and Montalvo 1994; Brown 2004; López 2008; Araújo and Borrell 2006).

Skin Color Stratification Within Various Ethnoracial Groups

Much of the aforementioned research aimed toward backing the increasing significance of skin color thesis, while convincing, tends to focus on intragroup disparities in macro-level outcomes–especially disparities found between light and dark-skinned blacks. Less of this work considers how skin color stratification influences adolescents specifically in their daily peer interactions. To be sure, past research on the effect of skin color has traditionally focused on the advantages light-skinned blacks hold relative to blacks of darker skin tones (Hughes and Hertel 1990). Over the years, however, literature on colorism has expanded to include empirical studies on the impact of skin color within other ethnoracial groups. For instance, within the realm of research on education, Murguia and Telles find that lighter-skinned Latino survey respondents completed more years of schooling net of differences in family background (1996). Additionally, a body of literature on panethnicity among Asians attests to the academic achievement gap between East Asians of lighter complexion (i.e., the Chinese, Japanese, and Koreans) and Southeast Asians of darker complexion (i.e., Laotians, Hmong, Cambodian and Vietnamese) (Kelly 1986; Zhou and Xiong 2005; Ngo and Lee 2007). Some research on criminal justice has also broadened the scope of the ethnoracial groups studied with respect to disparities by skin color. For example, Finkeldey and Demuth investigate intersecting effects of self-identity race/ethnicity on experiencing an arrest in adulthood ultimately finding that Latinos and Native Americans of lighter complexion are less likely to experience arrest than their darker-skinned counterparts (2021).

Skin Color Stratification Among Adolescents

Although intragroup disparities in broad macro-level outcomes represent a significant share of the social scientific literature on racial phenotype, few focus specifically on adolescents. Most of the existing research on adolescents tends to center on the association between racial appearance and overall well-being including mental and sexual health. For example, social psychologist, Patricia Louie, found that among a nationally representative sample of black adolescents, respondents with darkest skin tones exhibited higher levels of depressive symptoms than their lighter skin tone peers (2019). When considering proper psychological disorders indicative of severe impairment and dysfunction, the association shifted in that it showed significant differences between very dark and medium brown skin tones, but no significant differences when compared to those of very light complexion. Using data from a longitudinal study on 397 African American young women, Landor and collogues found that skin color was linked to sexual behavior and sexual health outcomes in the direction skin color hierarchies traditionally operate, however this association was buffered by high levels of parental support

(2019). Other research found results suggestive of a stronger desire for light-skinned dating partners among male adolescents (Roscoe, Diana and Brooks 1987).

Skin Color and Adolescent Friendships

Of all empirical studies on adolescents and skin color stratification, none to our knowledge explicitly consider the degree to which adolescent friendships cross skin color lines, thereby potentially masking overlooked racial segregation in voluntary friendship networks. The small number of studies that have aimed at investigating the effects of skin color on friendship networks, while illuminating, are limited in that they examine adult networks rather than those of adolescents and they either tend to focus on one ethnoracial group alone or do not investigate skin color homophily itself. To be sure, in their study of color homophily among Dominican and Puerto Rican immigrants, Roth and Marin examine egocentric networks ultimately finding evidence of homophily by skin tone among Dominican subjects (2021). Additionally, using data from the National Longitudinal Survey of Freshman and a series of multinomial logit models Santana finds lighter-skinned members of lower status ethnoracial groups are more likely to have close friendships with members of higher status ethnoracial groups (2022).

HYPOTHESIS 1: *Are adolescents similar in skin color more likely to be tied to one another within friendships, net of homophily based on racial self-classification?*

Both racial and skin color hierarchies have been shown to play a consequential role in life chances among individuals, and their effects are not mutually exclusive. Indeed, race scholars have argued dark-skin discrimination occurs within as well as across races (Turner 1995). Some evidence even suggests that intra-racial disparities are as detrimental to a person’s life chances as disparities traditionally associated with racial divisions (Hughes and Hertel 1990).

Therefore, because ethnoracial positionality in the social structure that is the U.S. racial hierarchy is defined in part by one's racial self-identification *and* one's racial appearance or phenotype, and ethnoracial positionality in turn influences who adolescents befriend, I hypothesize that adolescents of similar skin tone will be more likely to be tied to one another relative to those not similar in skin tone, net of racial self-classification homophily effects. In other words, I hypothesize that adolescents similar in skin color will be more likely to be tied to one another, even after controlling for homophily by race.

HYPOTHESIS 2: *Is there evidence that the extent of skin color homophily differs across ethnoracial groups?*

Considering evidence suggesting the increasing significance of skin color in determining individual group position coupled with evidence suggesting a broadening of the white racial category, I hypothesize that if there is evidence of skin color homophily within adolescent friendship networks, this association will be moderated by ethnoracial self-classification (i.e., ethnoracial identity). Specifically, I predict that Black adolescents will be more likely to engage in cross-race ties with outgroup members of similar skin tone due to the stringent racial boundary placing them at the bottom of the U.S. racial order and the subsequent dearth of white and white proximal networks available to them, which are characterized by individuals of lighter skin tones.

Data & Methods

To answer these theory-driven questions the present study uses primary data from the Teen Identity Development and Education Study (TIDES). This study involved survey data collection from the entire student body of two ethnically diverse high schools at three periods of time over a one-year period (N_{W1} = 3,191, N_{W2} = 3,605, N_{W3} = 3,109) in two distinct geographical regions of the U.S.– one in the Southwest and the other in the Midwest (see Table 5 for survey consent rate). In each survey, participants reported on a slew of topics, including demographic characteristics, friendship networks, experiences of ethnic/racial discrimination, sense of school belonging, and depressive symptoms to name a few.

Self-reported skin color data was not collected; instead, an innovative data collection design leveraging advances in computer software was developed to further reduce measurement error. To be sure, while self-perceived skin color is one of the most cost-effective ways of measuring skin tone, it is not ideal for investigation of skin color stratification since others' perceptions of are highly influential in determining that stratification. Additionally, interviewer classification struggles to reconcile the fact that perceptions are skewed by the individual doing the perceiving. For instance, Hill found greater variance in color measurement among Black and White interviewers when classifying respondents within their own race (2002), making the advantages of repeated measurement by multiple observers evident. Therefore, and akin to the strategy used by other scholars (Feliciano 2015; Feliciano & Robnett 2014), subjects' color yearbook headshots were classified by multiple observers (i.e., coders). The coding of subject skin tone involved a matching procedure wherein a skin color scale was displayed on the same computer screen as a subject headshot and coders selected the color on the scale that most resembled the skin tone of the individual in the photo. Every headshot was coded by no less than three coders. Coders were assigned a set of headshots to code at random.

Coding of subject headshots occurred over the course of four academic quarters. Coders were comprised of a total of 38 students, one of which was a graduate student. The group of coders was an ethnoracially diverse set of students, female undergraduate research assistants from ethnoracially diverse backgrounds, barring representation of black coders, as only one of the 38 coders identified as Black/African-American. The majority of coders were female, with 7 coders identifying as male and 1 identifying as being of some other gender. Coders were recruited from courses at the University of California, Irvine and all received course credit for their work. To help diminish the effects of biased ratings due to coder race, skin color will be operationalized as the average rating across all ratings.

Measures

In what follows, we describe the operational definitions for the outcome, predictor, and control variables.

Friendship

Friendship nominations were obtained via a survey questionnaire, wherein respondents were asked to list the "*friends you hang around with and talk to the most in your school.*" Respondents were instructed to list as many or as few names as they needed, but to be sure to list closest friends only. The TIDES procedure for collecting friendship nomination data differs from the procedure used in Add Health–the most-widely used adolescent friendship nomination data–in that it does not restrict nominations by gender. Rather than asking adolescents to name their five closest male friends and their five closest female friends adolescents were permitted to name as many friends as they wanted irrespective of their nominee's gender, thereby providing more reliable friendship data.

Skin color homophily

Skin color homophily uses the Project on Ethnicity and Race in Latin America (PERLA) color palette as the instrument for measuring skin color. The PERLA skin color measure is an 11-point scale where lower values are assigned to lighter colors and the measure was coded into three categories–light (1-3), medium (4-5), and dark (6+). Skin color homophily is specified as a “nodematch” term in the ERGM models. “Nodematch” terms account for the number of edges (i.e., friendship ties) whose incident nodes match on the value of a nodal attribute. Simply put, the skin color “nodematch” term accounts for whether the adolescents engaged in a friendship fall within the same skin color category.

Racial homophily

Racial homophily uses self-identified race from the TIDES survey. Although there are three waves of data, we assume race to be temporally invariant and use the race at first mention to racially categorize individuals. Racial categories include “Black/African American”, “Latino or Hispanic”, “Asian”, “American Indian/Native American”, “White”, “Other”, and a “missing” category. A racial self-classification “nodematch” term was specified that included a differential option. The differential option allows for the extent of homophily to differ across ethnoracial groups. Simply put, the racial homophily terms accounts for the extent to which adolescents engaged in a friendship fall within the same racial category, the extent of which may differ from group to group.

Controls

Edges: The edges term is akin to an intercept in regression analysis. In other words, the edges term accounts for the likelihood of a friendship tie being present in a given network.

Other forms of homophily: Due to the fact that homophily based on other attributes are sure to be present in the networks under investigation, we control for some of the theoretically most salient forms of homophily. These include homophily based on gender, grade, and highest level of parents’ completed education.

Nodal attributes: Models also control for adolescent skin color and adolescent racial self-identification since these attributes may influence the likelihood of an individual nominating a friend or being nominated by a friend and because excluding them can lead to misinterpretation of the skin color homophily effect.

Edge attributes: We include a control for whether an edge (i.e., friendship tie) exists between two nodes (i.e., adolescents) that share the same extracurricular activity due to the fact that shared interests and additional time spent in proximity outside the formal classroom likely increase the odds of a friendship forming.

Network structure: Models attempt to control for endogenous structural characteristics of the networks including mutuality, transitivity, open triads, indegree and outdegree. Mutuality measures the likelihood of a friendship tie being reciprocal between two individuals in a given network and transitivity measures the likelihood of an individual's two friends being friends themselves. Open triads account for number of shared friends, regardless of whether those shared friends are friends themselves. Indegree and outdegree terms account for the geometrically weighted incoming tie and outgoing tie distribution of a given network.

Analytic Strategy

To answer the question of whether adolescents similar in skin tone are more likely to be tied to one another relative to those dissimilar in skin color, the proposed study will use a Markov chain Monte Carlo (MCMC) technique to simulate random graphs (i.e., networks) from a specified exponential random graph model (ERGM). ERGMs were chosen to investigate the relationship between skin tone stratification and adolescent friendship networks because of their unique properties. To illustrate, ERGMs are stochastic models, which allow us to capture both the regularities in the process giving rise to network ties, while also recognizing that there is variability that is unlikely to be modeled in detail. This stochasticity is important when the process being modeled is in effect stochastic in nature. These statistical models also allow for inference about whether certain network structures are more typical in specific models than might be expected by chance, thereby allowing us to distinguish between endogenous structural effects (i.e., structural balance) and node-level effects (i.e., homophily) and to distinguish between their respective contributions (Robins et al. 2007). Finally, ERG models use relatively new procedures to evaluate how well a model fits the observed network by comparing structural statistics of the observed network to the corresponding statistics on networks simulated from the fitted model (Hunter, Goodreau and Handcock 2012). Together these unique attributes make this exponential family of models well-suited for the research questions at hand.

The full TIDES dataset contains two schools, measured across three waves and ERGMs will be specified for each school wave separately. The edges (i.e., friendship ties) in these raw network data are directed, because it is possible for subject A to name subject B as a friend without subject B naming A. This analysis therefore considers the likelihood of being nominated by someone as a friend *or* nominating someone as a friend, but those nominations need not be reciprocated.

Friendship networks will be represented by a symmetric n x n matrix $\boldsymbol{Y}$ and a n x q matrix $\mathbf{X}$ of nodal covariates, where n is the number of nodes (or actors). The cells within the adjacency matrix $\mathbf{Y}$ contain 0's and 1's, with $Y_{ij} = 1$ if the presence of a tie between i and j exists and $Y_{ij} = 0$ when it does not. Because self-nominations were not permitted, the diagonal of the adjacency matrix is set to 0 for all i.

The nodal covariate matrix $\mathbf{X}$ may include many measurements of individual attributes for i, including both exogenous (i.e., age, gender, racial self-classification, skin color, etc.) and endogenous ones (tobacco use attribute influenced by friendships). This analysis however will only include attributes which are assumed to be fixed and exogenous.

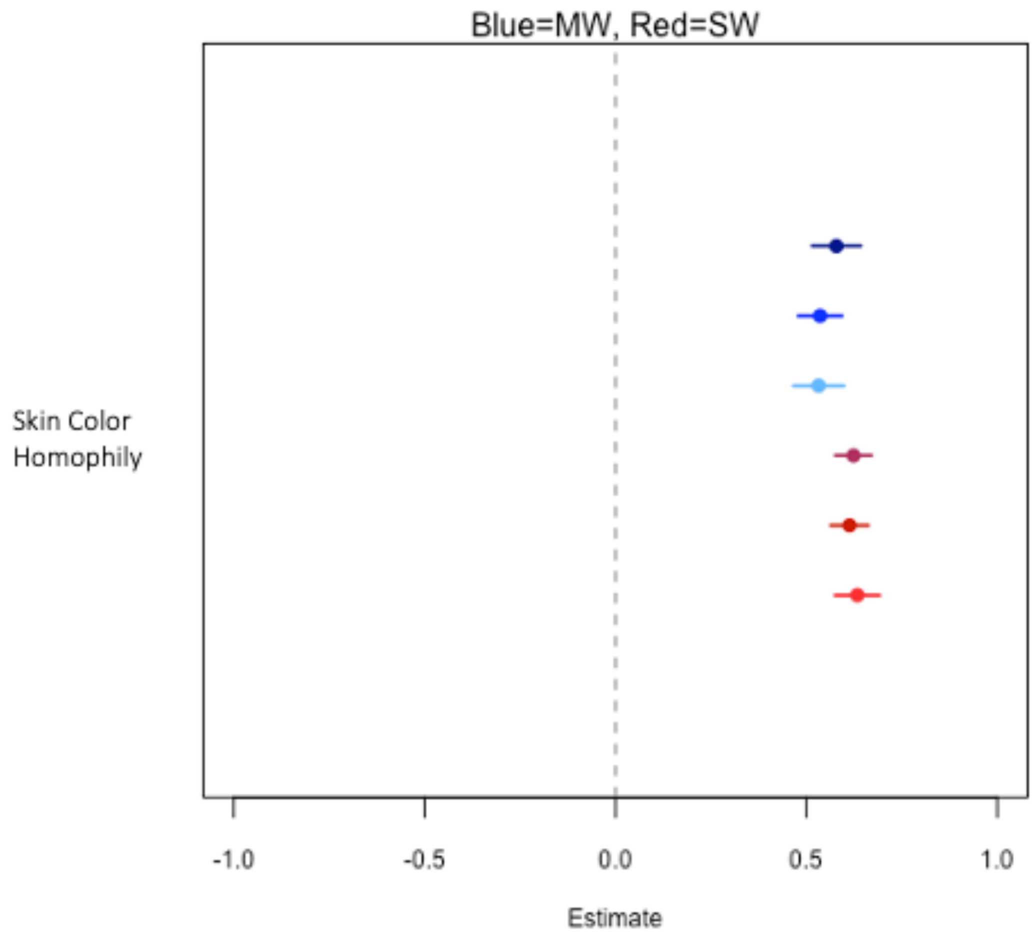
To reiterate, the overall aim in using ERGMs is to model the random behavior of the adjacency matrix **Y** conditional on the covariate matrix **X** however not every covariate is included in every model. Instead, models are built through a stepwise process. The first set of models represent the baseline model which includes an edges term and a skin color homophily term. The second set of models is the baseline model with the addition of a racial homophily term. The third set of models includes skin color homophily and racial homophily as key predictors, in addition to the set of controls (i.e., other forms of homophily, nodal attributes, and an edge attribute). Finally, the fourth set of models include skin color homophily, racial homophily, controls with the addition of an interaction effects between skin color homophily and racial self-identification as a nodal attribute.

Results

Skin Color Homophily in Adolescent Friendships

Figure 4 plots the skin color homophily coefficients from the ERGMs for each school by wave. The three blue points denote Midwest skin color homophily estimates and the three red points denote those from the Southwest school–one point for each wave of network data by school. The colored lines running through the center of the plotted points denote the 95% confidence intervals for each estimate. All plotted skin color homophily estimates are positive, with slightly higher estimates in the Southwest region and small confidence intervals. Tables 6-11 reveal that each of the skin color homophily estimates are significant at the $p<0.001$ level. Calculated odds ratios range from 1.70-1.78 and 1.85-1.88 in the Midwest and Southwest regions, respectively. Odds ratios stemming from homophily estimates are interpreted as the likelihood of a tie being present between two individuals given similarity on a certain attribute relative to the likelihood of that tie being present at random. Thus, in the Midwest school the likelihood of a friendship being present between two adolescents with similar skin tone is, on average, 1.73 times higher than the likelihood of the friendship being present at random. Slightly higher estimates in the Southwest school reveal that, on average, adolescents are 1.87 times more likely to befriend or be befriended by an individual of similar skin tone, than they are to befriend or be befriended by a peer in their school at random. The association between similarity on skin color and the likelihood of a friendship tie existing is not surprising, as the baseline model has yet to control for racial homophily and therefore skin color homophily may simply be serving as a proxy for racial homophily since skin color and racial self-identification tend to be highly correlated.

Figure 4. Baseline ERGMs' plotted coefficients by school and wave

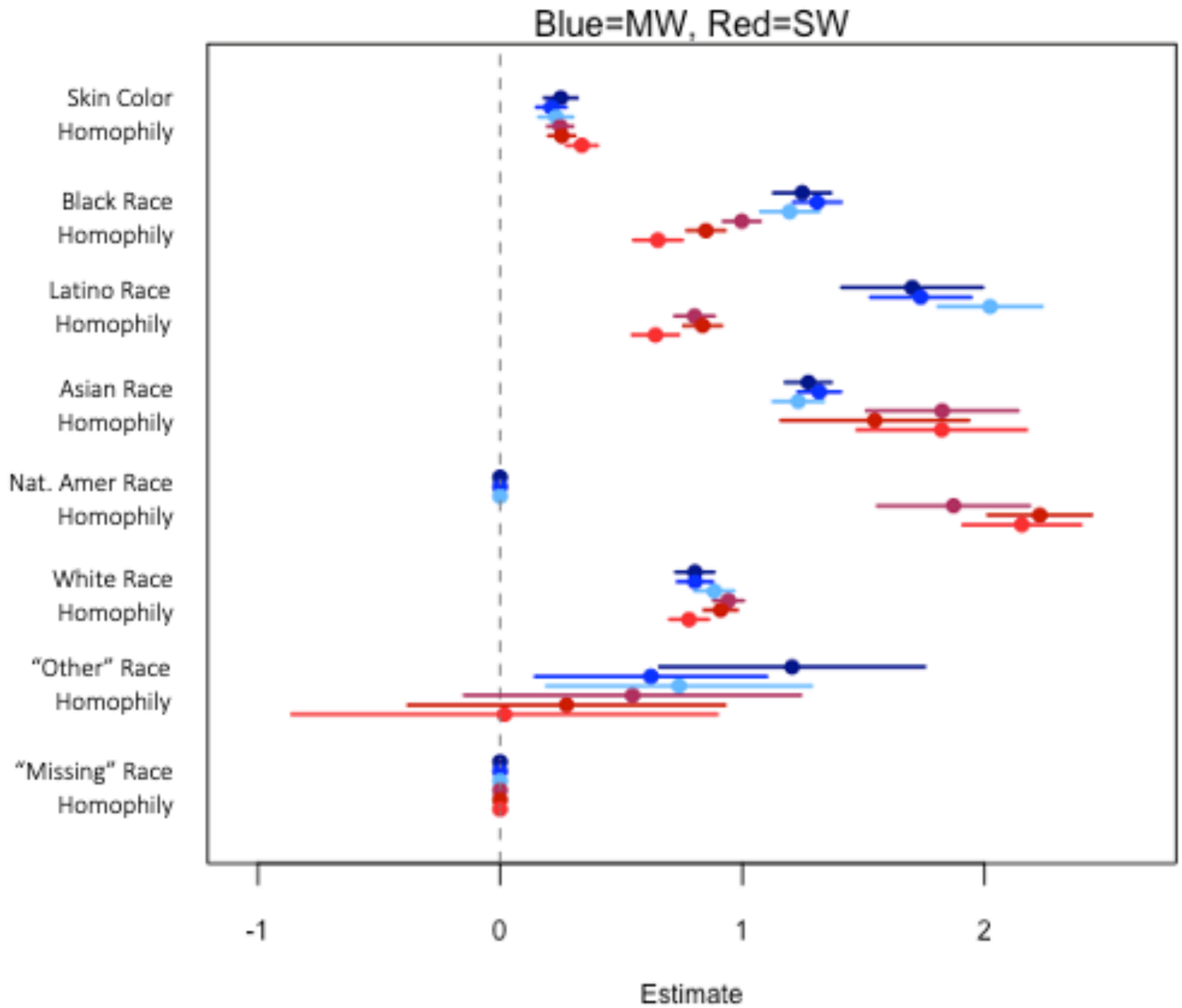


Skin Color Homophily in Adolescent Friendships Net of Racial Homophily

Figure 5 contains the plotted estimates for the covariates included in Models 2 by school and wave. Again, Models 2 include an edges term (estimate for edges term not plotted in figure), a skin color homophily term, and a racial homophily term. The first set if plotted coefficients correspond to the skin color homophily estimates. The remaining set of plotted coefficients correspond to the estimates of racial homophily–one for each of the ethnoracial categories captured in our measure of race. Notably, the skin color homophily estimates are still positive and their confidence intervals do not overlap with zero, meaning skin color homophily results are robust to the inclusion of racial homophily in Model 2 for each school across all three waves. Substantively, results reveal that adolescents are more likely to have cross-race friendships with peers of similar skin tone than they are to have cross-race friendships with adolescents of different skin tone.

Racial homophily, in accordance with previous research, is stronger relative to homophily by skin tone for every ethnoracial group, as made evident by higher estimates, barring racial homophily among Native Americans in the Midwest likely due to their extremely small sample size (N<14 across all waves). Native Americans in the Southwest sample, however, display the strongest tendency to have friends of similar ethnoracial background. Latino subjects in the Midwest sample and Asian subjects in the Southwest sample produce the second highest racial homophily estimates, likely due their small sample size in those respective schools and the tendency for smaller racial groups to be more insular. Black adolescents have more racially homophilous friendships in the Midwest school than they have in the Southwest school. Finally, white adolescents' tendency to have friends of the same race is similar across both schools and is akin to that of Latino and Black adolescents in the Southwest.

Figure 5. Model 2 ERGMs' plotted coefficients by school and wave



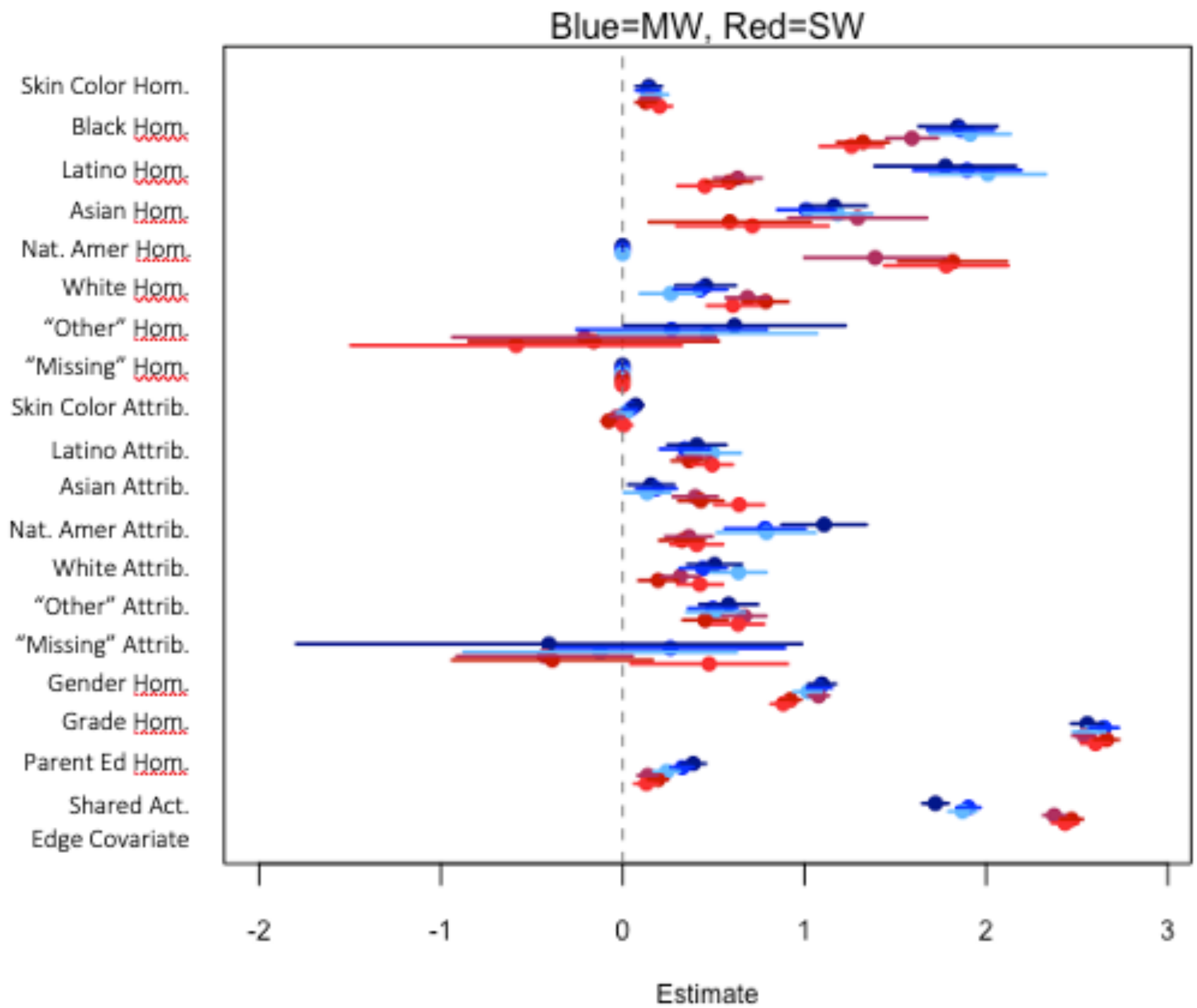
Skin Color Homophily in Adolescent Friendships Net of Racial Homophily and Other Controls

Plotted coefficients from Model 3 reveal that skin color homophily is a significant predictor of a friendship tie existing between two adolescents and that this effect is robust to the inclusion of racial homophily, homophily of other attributes which have been shown to breed connection, nodal covariates, and an edge covariate that denotes shared extracurricular activity. That most of the skin color nodal covariate estimates are plotted at 0 reveal little to no significance for skin color determining the degree to which adolescents have friends or are befriended by others. To be sure, Tables 6 and 7 show the results for the Midwest school in Wave 1 and Wave 2, respectively, and contain the only significant positive nodal skin color estimates within Model 3. That these significant associations are positive reveals that individuals of darker skin tone tend to have more friendship ties in general, but only for this school at these two waves. The race nodal covariates reveal that race is a significant predictor of the likelihood of tie being present between two adolescents but whether race is a stronger predictor of a friendships for one ethnoracial group over another does not appear to be evident, given that the estimate confidence intervals overlap.

Evidence of a significant gender homophily and grade homophily effect is also shown by their respective plotted coefficients. However, in line with evidence on the effects of propinquity on tie formation, grade homophily is a far stronger predictor of friendship than similarity based on gender. Serving as a proxy for homophily based on socioeconomic status, homophily based on similar parental level of education is significant across both schools and all three waves. That the addition of homophily of parent level of education does not eliminate the effect of similarity based on skin tone suggests skin color homophily is not simply an artifact of adolescents tending to be friends with individuals of the same socioeconomic background –a point of substantive interest since skin tone and socioeconomic background have been shown to be highly correlated. Whether friendships are more likely to exist between adolescents sharing an extracurricular activity is also made evident by however the effect appears to be stronger in the Southwest school.

After adding these set of controls, homophily based on race continues to be a significant predictor of friendship, however changes in the sizes of the coefficients are particularly marked among Black subjects, where we find consistent increases across both schools and all three waves (see Model 3 in Tables 6-11).

Figure 6. Model 3 ERGMs' plotted coefficients by school and wave

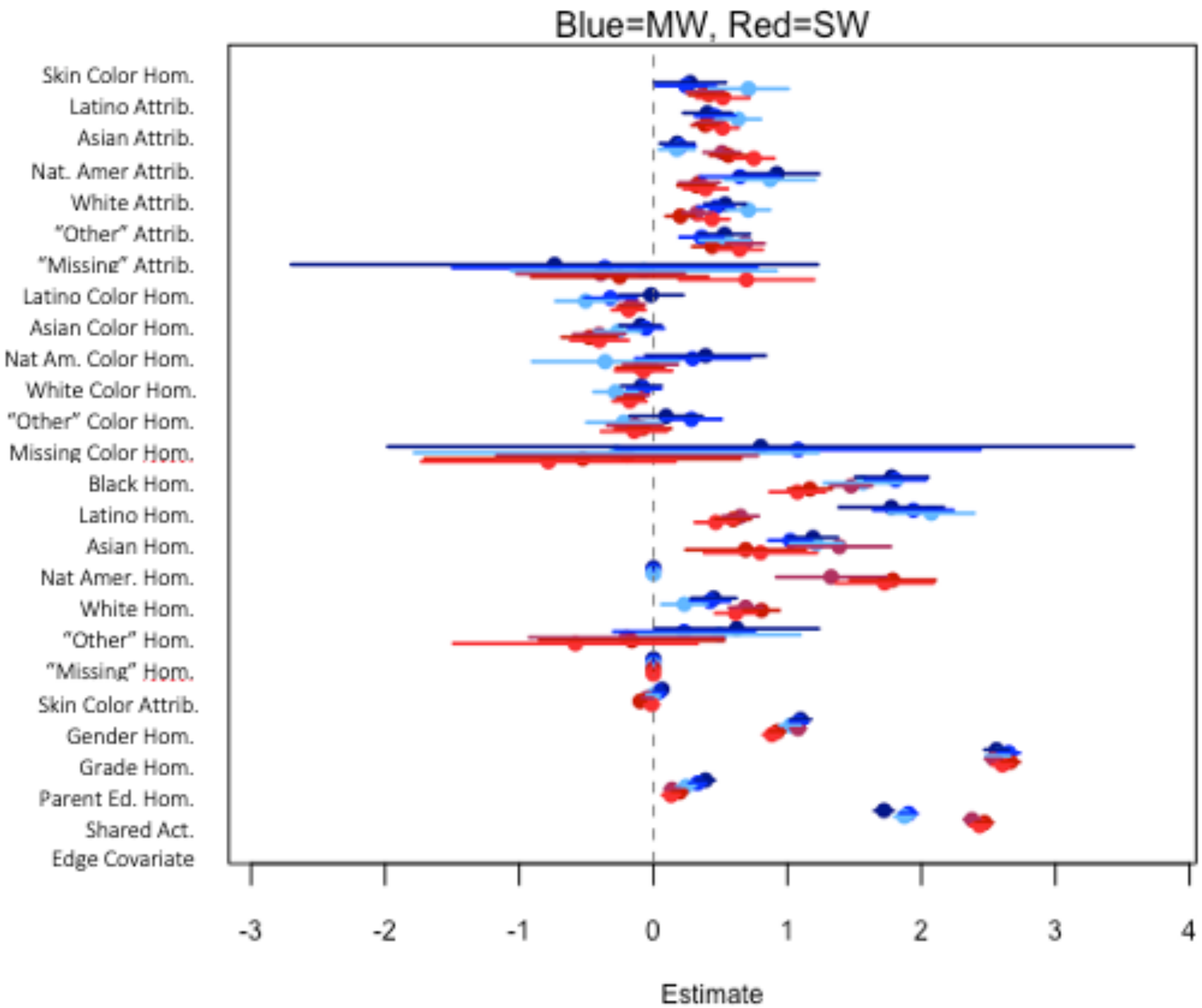


Differences in Skin Color Homophily by Race?

Figure 7 shows plotted coefficients for Models 4 found in Tables 6-11. Model 4 includes skin color homophily and racial homophily as key predictors and all the controls present in Models 3, with the addition of interacting effects between skin color homophily and race as a nodal attribute. While most interaction coefficients in Figure 7 are negative, Tables 6-11 show not all are significant. For example, none of the interacting effects are significant among the Midwest Wave 1 sample (see Table 6), yet on average consistent results emerge as the analysis is conducted across school/waves. To be sure, statistically significant interaction effects between homophily by skin tone and identifying as Latino ($p<.001$ level) are shown in Models 4 among the Midwest Waves 2 and 3 samples (see Tables 7 and 8) and in the Southwest sample across Waves 1-3 (see Tables 9-11). The Asian sample and the White sample also show statistically significant negative interacting effects with skin color homophily in the Midwest school at Wave 3 (see Table 8) and in the Southwest school across Waves 1-3 (see Tables 9-11). That each of the significant interacting effects are negative, relative to Black subjects suggests that Black adolescents are more likely to have ties to cross race friends similar in skin tone than their Latino, Asian and White peers.

Additional calculations producing interacting effect confidence intervals reveal that in the Southwest school across all waves, not merely are Asians adolescents less likely than Black subject to have homophilous ties by skin tone, they are not likely to have homophilous skin color ties at all, as made evident by respective confidence intervals overlapping zero (see Table 12). Similar findings are evident among Latino subjects in the Midwest Waves 2 and 3 sample, however, not in any of the Southwest sample waves. To be sure, while Latinos in the Midwest sample, on average, appear to not exhibit a tendency to form skin color homophilous ties with cross-race friends, Latinos in the Southwest sample do tend to have homophilous skin color ties to cross-race friends, however that tendency is weaker than it is for Black subjects. White subjects' tendency to have cross-race friendship ties with those similar in skin tone is the same in both the Midwest and Southwest school across all waves, however that association one average is also weaker, relative to the association found among their black peers.

Figure 7. Model 4 ERGMs' plotted coefficients by school and wave



Discussion

Overall, findings from the ERGM analysis reveal that similarity in skin tone structures adolescent friendship networks by influencing the likelihood of a friendship existing between two individuals, over and above homophily by race, gender, grade, and socioeconomic status, and a set of other meaningful controls. That skin color homophily cannot be explained away by racial homophily in any of our models shows that former racial exclusionary network studies' unidimensional operationalization of race has limited the detectability of the extent of racial segregation within friendship networks.

Using friendship nomination data collected at two ethnoracially diverse schools at three different points in time I examined the tendency toward voluntary social connectivity with similar others on the basis of skin color. Results affirm that similarity in skin tone was associated with an increased likelihood of a friendship tie existing between adolescents. In addition to confirming prior literature on the tendency toward befriending or being befriended by those from one's own ethnoracial background (Shrum, Cheek, and Hunter 1988; Joyner and Kao 2000; Moody 2001; Goodreau, Kitts and Morris 2009), these results demonstrate that what had previously been observed as heterogeneous friendship ties are, in a statistically significant number of cases, in fact racially homogenous.

We observed ethnoracial differences in the increased likelihood of cross-race ties existing between individuals of similar skin tone. That black adolescents, on average, exhibited the largest extent of skin color homophilous networks, relative to their Latino, Asian, and White counterparts underscores the relevance of sociological theory on prejudice and group position. To be clear, Blumer's work on racial attitudes which posits that such attitudes should be conceptualized as general orientation containing normative ideas of where one's own group should stand relative to outgroup members and these normative ideas, in turn, influence whom and how individuals interact with one another. That black individuals are least able to cross skin color boundaries in their cross-race ties, speaks to the racial rigidity they must operate under. Simply put, whether race is operationalized in the traditional sense (i.e., racial self-classification) or by its ocular nature (i.e., phenotypic appearance), black individuals are fixed within the group position they fall under.

While former studies have laid a strong foundation for the investigation of racial exclusionary networks, neglecting to consider multiple dimensions of race, particularly racial appearance, has resulted in potentially conservative estimates of meaningful racial segregation–segregation that has the ability to impede the salutary effects diverse social networks produce.

Paper 2: References

Allen, Walter, Edward Telles, and Margaret Hunter. 2000. "Skin Color, Income and Education: A Comparison of African Americans and Mexican Americans." *National Journal of Sociology* 12(1):129–80.

Anderson, Claud and Rue L. Cromwell. 1977. "'Black Is Beautiful' and the Color Preferences of Afro-American Youth." *The Journal of Negro Education* 46(1):76–88.

Araújo, Beverly Y. and Luisa N. Borrell. 2006. "Understanding the Link Between Discrimination, Mental Health Outcomes, and Life Chances Among Latinos." *Hispanic Journal of Behavioral Sciences* 28(2):245–66.

Blair, Irene V., Charles M. Judd, and Kristine M. Chapleau. 2004. "The Influence of Afrocentric Facial Features in Criminal Sentencing." *Psychological Science* 15(10):674–79.

Blau, Peter Michael. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure.* New York: Free Press.

Blumer, Herbert. 1958. "Race Prejudice as a Sense of Group Position." *The Pacific Sociological Review* 1(1):3–7.

Bobo, Lawrence D. 1999. "Prejudice as Group Position: Microfoundations of a Sociological Approach to Racism and Race Relations." *Journal of Social Issues* 55(3):445–72.

Bonilla-Silva, Eduardo. 2002. "Where Is the Love? A Rejoinder by Bonilla-Silva on the Latin Americanization Thesis." *Race and Society* 5(1):103–14.

Brown, Kendrick T. 2004. "The Power of Perception: Skin Tone Bias and Psychological Well-Being for Black Americans." Pp. 111–23 in *Racial identity in context: The legacy of Kenneth B. Clark.*, edited by G. Philogène. Washington: American Psychological Association.

Burch, Traci. 2015. “Skin Color and the Criminal Justice System: Beyond Black-white Disparities in Sentencing.” *Journal of Empirical Legal Studies* 12(3):395–420.

Cartwright, Dorwin. 1979. “Contemporary Social Psychology in Historical Perspective.” *Social Psychology Quarterly* 42(1):82-93.

Codina, G. Edward and Frank F. Montalvo. 1994. “Chicano Phenotype and Depression.” *Hispanic Journal of Behavioral Sciences* 16(3):296–306.

Coser, Lewis A. 1971. *Masters of Sociological Thought: Ideas in Historical and Social Context*. New York: Harcourt Brace Jovanovich.

Denton, Nancy A., and Douglas S. Massey. 1989. “Racial Identity Among Caribbean Hispanics: The Effect of Double Minority Status on Residential Segregation.” *American Sociological Review* 54(5):790.

Durkheim, Émile. 1893. “De la Division du Travail Social: Étude sur L’Organisation des Sociétés Supérieures.” Alcan, Paris.

Edwards, Korie, Katrina Carter-Tellison, and Cedric Herring. 2004. “For Richer, for Poorer, Whether Dark or Light: Skin Tone, Marital Status, and Spouse’s Earnings.” Pp. 65–81 in *Skin deep: How race and complexion matter in the “color-blind” era*, edited by C. Herring, V. Keith, and H.D. Horton. Chicago, Illinois: University of Illinois Press.

Feliciano, Cynthia and Belinda Robnett. 2014. “How External Racial Classifications Shape Latino Dating Choices.” *Du Bois Review: Social Science Research on Race* 11(2):295–328.

Finkeldey, Jessica G., and Stephen Demuth. 2021. “Race/Ethnicity, Perceived Skin Color, and the Likelihood of Adult Arrest.” *Race and Justice* 11(4):567–91.

Goodreau, Steven M., James A. Kitts, and Martina Morris. 2009. “Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks*.” *Demography* 46(1):103–25.

Hakken, Jon. 1979. Discrimination against Chicanos in the Dallas Rental Housing Market: An Experimental Extension of the Housing Market Practices Survey. Division of Evaluation, US Dept. of Housing and Urban Development, Office of Policy Development and Research.

Helg Aline. 1990. ‘Race in Argentina and Cuba, 1880–1930: Theory, Policies, and Popular Reaction.” Pp. 37-69 in *The Idea of Race in Latin America, 1870–1940.*, edited by R. Graham. Austin University of Texas Press.

Hughes, Michael and Bradley R. Hertel. 1990. “The Significance of Color Remains: A Study of Life Chances, Mate Selection, and Ethnic Consciousness among Black Americans.” *Social Forces* 68(4):1105–20.

Hunter, Margaret. 2016. “Colorism in the Classroom: How Skin Tone Stratifies African American and Latina/o Students.” *Theory Into Practice* 55(1):54–61.

Hunter, David R., Steven M. Goodreau, and Mark S. Handcock. 2012. “Goodness of Fit of Social Network Models.” *Journal of the American Statistical* Association 103(481):248–58.

Ibarra, Herminia. 1995. “Race, Opportunity, and Diversity of Social Circles in Managerial Networks.” *Academy of Management Journal* 38(3):673–703.

Johnson, Doyle Paul. 1981. *Sociological Theory: Classical Founders and Contemporary Perspectives*. New York: Wiley.

Joyner, Kara and Grace Kao. 2000. “School Racial Composition and Adolescent Racial Homophily.” *Social Science Quarterly* 810–25.

Kalmijn, Matthijs. 1998. “Intermarriage and Homogamy: Causes, Patterns, Trends.” *Annual Review of Sociology* 24(1):395–421.

Kao, Grace and Kara Joyner. 2004. “Do Race and Ethnicity Matter among Friends? Activities among Interracial, Interethnic, and Intraethnic Adolescent Friends.” *The Sociological Quarterly* 45(3):557–73.

Kelly, Gail P. 1986. “Coping with America: Refugees from Vietnam, Cambodia, and Laois in the 1970s and 1980s.” *The ANNALS of the American Academy of Political and Social Science* 287(1):138-149

Kreisman, Daniel and Marcos A. Rangel. 2015. “On the Blurring of the Color Line: Wages and Employment for Black Males of Different Skin Tones.” *Review of Economics and Statistics* 97(1):1–13.

Landor, Antoinette M., Leslie Gordon Simons, Ellen M. Granberg, and Janet N. Melby. 2019. “Colorizing Self-Esteem among African American Young Women: Linking Skin Tone, Parental Support, and Sexual Health.” *Journal of Child and Family Studies* 28(7):1886–98.

Lincoln, James R., and Jon Miller. 1979. “Work and Friendship Ties in Organizations: A Comparative Analysis of Relation Networks.” *Administrative Science Quarterly* 24(2):181–99.

Louie, Patricia. 2020. “Revisiting the Cost of Skin Color: Discrimination, Mastery, and Mental Health among Black Adolescents.” *Society and Mental Health* 10(1):1–19.

Marsden, Peter V. 1987. “Core Discussion Networks of Americans.” *American Sociological Review* 52(1):122–31.

Marsden, Peter V. 1988. “Homogeneity in Confiding Relations.” *Social Networks* 10(1):57–76.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. “Birds of a Feather: Homophily in Social Networks.” *Annual Review of Sociology* 27(1):415–44.

Mead, George Herbert. 1934. *Mind, Self & Society from the Standpoint of a Social Behaviorist*., edited by Morris, C. W. Chicago, Ill.: University of Chicago Press.

Murguia, Edward, and Edward E. Telles. 1996. “Phenotype and Schooling among Mexican Americans.” *Sociology of Education* 69(4):276.

Ngo, Bic, and Stacey J. Lee. 2007. “Complicating the Image of Model Minority Success: A Review of Southeast Asian American Education.” *Review of Educational Research* 77(4):415-453.

Portes, Alejandro, and Rubén G. Rumbaut. 2014. *Immigrant America: A Portrait*. 4th edition. Berkeley, California: University of California Press.

Robins, Garry, Pip Pattison, Yuval Kalish, and Dean Lusher. 2007. “An Introduction to Exponential Random Graph (P*) Models for Social Networks.” *Special Section: Advances in Exponential Random Graph (P*) Models* 29(2):173–91.

Roediger, David R. 1999. *The Wages of Whiteness: Race and the Making of the American Working Class*. London; New York: Verso.

Roscoe, Bruce, Mark S. Diana, and Richard H. Brooks. 1987. “Early, Middle, and Late Adolescents’ Views on Dating and Factors Influencing Partner Selection.” *Adolescence* 22(85):59.

Roth, Wendy. 2012. *Race Migrations: Latinos and the Cultural Transformation of Race*. Stanford, California: Stanford University Press.

Roth, Wendy D., and Alexandra Marin. 2021. "The Role of Skin Color in Latino Social Networks: Color Homophily in Sending and Receiving Societies." *Sociology of Race and Ethnicity* 7(2):175–93.

Santana, Emilce. 2022. "FRIENDSHIP IS SKIN (COLOR) DEEP: The Role of Skin Color in Cross-Ethnoracial Friendships." *Du Bois Review: Social Science Research on Race* 19(1):147–73.

Shrum, Wesley, Neil H. Cheek, and Saundra MacD. Hunter. 1988. "Friendship in School: Gender and Racial Homophily." *Sociology of Education* 61(4):227–39.

Waters, Mary C. 1990. *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.

Weber, Max. 1958. *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribner.

Wimmer, Andreas and Kevin Lewis. 2010. "Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook." *American Journal of Sociology* 116(2):583–642.

Yinger, John. 1991. *Housing Discrimination Study: Incidence of Discrimination and Variation in Discriminatory Behavior*. US Dept. of Housing and Urban Development, Office of Policy Development and Research.

Zhou, Min and Yang Sao Xiong. 2005. "The Multifaceted American Experiences of the Children of Asian Immigrants: Lessons for Segmented Assimilation." *Ethnic and Racial Studies* 28(6):1119–52.

Table 5. Midwest and Southwest combined survey consent rate, school record rate, and opt out rate

	Wave 1 | Spring 2017	Wave 2 | Fall 2017	Wave 3 | Spring 2018
Survey Participation Rate $(\frac{surveys\ completed}{total\ \#\ students})$	82.00% (N = 3191)	86.94% (N = 3623)	79.18% (N = 3127)
Survey Opt Out Rate	.43% (N = 17)	1.27% (N = 53)	4.58% (N = 181)
Active Consent Return Rate $\frac{\#\ gave\ anything\ back}{total\ \#\ students}$	47.24% (N = 1862)	46.46% (N = 1939)	-
School Records Access Rate $\frac{\#\ approved}{total\ \#\ students}$	36.94% (N = 1456)	36.42% (N = 1520)	-
School Records Refusal Rate $\frac{\#\ refused}{total\ \#\ students}$	10.30% (N = 406)	10.04% (N = 419)	-
Roster for Active Consent	N = 3941	N = 4173	-
Roster for Survey Collection	N = 3891	N = 4167	N = 3949

Source: Teen Identity Development and Education Study

Table 6. Midwest Results of the Exponential Random Graph Models of Friendship at Wave 1

	Model 1			Model 2			Model 3			Model 4		
Effects	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***
Edges	-6.06	(0.02)	***	-6.35	(0.03)	***	-9.82	(0.16)	***	-9.81	(0.16)	***
Homophily												
Skin color (nodematch)	0.58	(0.03)	***	0.25	(0.03)	***	0.15	(0.04)	***	0.27	(0.14)	*
Gender (nodematch)							1.10	(0.04)	***	1.10	(0.04)	***
Grade (nodematch)							2.56	(0.04)	***	2.56	(0.04)	***
Parents' completed education (nodematch)							0.39	(0.03)	***	0.39	(0.03)	***
Race (nodematch, diff)												
Black				1.25	(0.06)	***	1.85	(0.11)	***	1.78	(0.14)	***
Latino				1.70	(0.15)	***	1.78	(0.20)	***	1.77	(0.20)	***
Asian				1.27	(0.05)	***	1.16	(0.09)	***	1.19	(0.09)	***
Native American				–	–	–	–	–	–	–	–	–
White				0.80	(0.04)	***	0.46	(0.08)	***	0.45	(0.08)	***
Other				1.21	(0.28)	***	0.62	(0.31)	*	0.62	(0.31)	*
Missing				–	–	–	–	–	–	–	–	–
Nodal Attributes												
Skin color							0.07	(0.02)	***	0.06	(0.04)	0.15
Race (ref: Black)												
Latino							0.41	(0.08)	***	0.40	(0.09)	***
Asian							0.16	(0.06)	**	0.18	(0.07)	***
Native American							1.11	(0.12)	***	0.92	(0.16)	***
White							0.51	(0.08)	***	0.53	(0.08)	***
Other							0.58	(0.08)	***	0.53	(0.10)	***
Missing							-0.40	(0.71)		-0.74	(1.00)	
Edge Attributes												
Activity Comembership							1.72	(0.04)	***	1.72	(0.04)	***
Network Structure												
Mutuality							–	–	–			
Transitivity							–	–	–			
Open triads												
Indegree (gw)							–	–	–			
Outdegree (gw)							–	–	–			
Skin Color Homophily x Ego Race												
Black (ref)												
Latino										-0.02	(0.12)	
Asian										-0.10	(0.08)	
Native American										0.39	(0.23)	†
White										-0.09	(0.08)	
Other										0.09	(0.14)	
Missing										0.80	(1.42)	
BIC	**51620**			**50846**			**41870**			**41945**		

Source: Teen Identity Development and Education Study. Notes: * p<0.05; ** p<0.01; *** p<0.001.

Table 7. Midwest Results of the Exponential Random Graph Models of Friendship at Wave 2

	Model 1			Model 2			Model 3			Model 4		
Effects	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***
Edges	-6.23	(0.02)	***	-6.52	(0.02)	***	-9.77	(0.15)	***	-9.77	(0.15)	***
Homophily												
Skin color (nodematch)	0.54	(0.03)	***	0.21	(0.03)	***	0.14	(0.03)	***	0.24	(0.11)	*
Gender (nodematch)							1.08	(0.03)	***	1.08	(0.03)	***
Grade (nodematch)							2.65	(0.04)	***	2.65	(0.04)	***
Parents' completed education (nodematch)							0.33	(0.03)	***	0.33	(0.03)	***
Race (nodematch, diff)												
Black				1.31	(0.05)	***	1.86	(0.09)	***	1.81	(0.12)	***
Latino				1.74	(0.11)	***	1.90	(0.15)	***	1.94	(0.15)	***
Asian				1.32	(0.04)	***	1.01	(0.08)	***	1.02	(0.08)	***
Native American				–	–	–	–	–	–	–	–	–
White				0.81	(0.04)	***	0.43	(0.07)	***	0.42	(0.07)	***
Other				0.62	(0.24)	**	0.27	(0.27)		0.23	(0.27)	
Missing				–	–	–	–	–	–	–	–	–
Nodal Attributes												
Skin color							0.04	(0.02)	*	0.04	(0.02)	†
Race (ref: Black)												
Latino							0.34	(0.07)	***	0.46	(0.08)	***
Asian							0.19	(0.06)	***	0.20	(0.06)	***
Native American							0.78	(0.11)	***	0.65	(0.15)	***
White							0.44	(0.07)	***	0.47	(0.07)	***
Other							0.50	(0.07)	***	0.36	(0.08)	***
Missing							0.26	(0.32)		-0.36	(0.58)	
Edge Attributes												
Activity Comembership							1.90	(0.03)	***	1.90	(0.03)	***
Network Structure												
Mutuality							–	–	–			
Transitivity							–	–	–			
Open triads												
Indegree (gw)							–	–	–			
Outdegree (gw)							–	–	–			
Skin Color Homophily x Ego Race												
Black (ref)												
Latino										-0.32	(0.10)	***
Asian										-0.06	(0.07)	
Native American										0.29	(0.22)	
White										-0.07	(0.06)	
Other										0.28	(0.11)	**
Missing										1.08	(0.69)	
BIC		**65812**			**64794**			**53062**			**53118**	

Source: Teen Identity Development and Education Study. Notes: * p<0.05; ** p<0.01; *** p<0.001.

Table 8. Midwest Results of the Exponential Random Graph Models of Friendship at Wave 3

	Model 1			Model 2			Model 3			Model 4		
Effects	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***
Edges	-6.37	(0.02)	***	-6.67	(0.03)	***	-9.81	(0.17)	***	-9.85	(0.17)	***
Homophily												
Skin color (nodematch)	0.53	(0.03)	***	0.23	(0.04)	***	0.18	(0.04)	***	0.71	(0.15)	***
Gender (nodematch)							1.02	(0.04)	***	1.02	(0.04)	***
Grade (nodematch)							2.57	(0.05)	***	2.57	(0.05)	***
Parents' completed education (nodematch)							0.24	(0.04)	***	0.24	(0.04)	***
Race (nodematch, diff)												
Black				1.20	(0.06)	***	1.91	(0.11)	***	1.56	(0.15)	***
Latino				2.02	(0.11)	***	2.01	(0.16)	***	2.07	(0.16)	***
Asian				1.23	(0.05)	***	1.18	(0.10)	***	1.22	(0.10)	***
Native American				–	–	–	–	–	–	–	–	–
White				0.88	(0.04)	***	0.26	(0.09)	***	0.23	(0.09)	***
Other				0.74	(0.28)	***	0.47	(0.31)		0.49	(0.31)	
Missing				–	–	–	–	–	–	–	–	–
Nodal Attributes												
Skin color							0.01	(0.02)		0.00	(0.02)	
Race (ref: Black)												
Latino							0.50	(0.08)	***	0.63	(0.09)	***
Asian							0.14	(0.06)	*	0.17	(0.07)	***
Native American							0.79	(0.14)	***	0.87	(0.17)	***
White							0.64	(0.08)	***	0.71	(0.08)	***
Other							0.51	(0.08)	***	0.53	(0.10)	***
Missing							-0.12	(0.38)		-0.07	(0.50)	
Edge Attributes												
Activity Comembership							1.87	(0.04)	***	1.87	(0.04)	***
Network Structure												
Mutuality							–	–	–			
Transitivity							–	–	–			
Open triads												
Indegree (gw)							–	–	–			
Outdegree (gw)							–	–	–			
Skin Color Homophily x Ego Race												
Black (ref)												
Latino										-0.51	(0.11)	***
Asian										-0.26	(0.09)	***
Native American										-0.36	(0.28)	
White										-0.28	(0.08)	***
Other										-0.22	(0.14)	
Missing										-0.28	(0.77)	
BIC	**50509**			**49795**			**41465**			**41529**		

Source: Teen Identity Development and Education Study. Notes: * p<0.05; ** p<0.01; *** p<0.001.

Table 9. Southwest Results of the Exponential Random Graph Models of Friendship at Wave 1

	Model 1			Model 2			Model 3			Model 4		
Effects	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***
Edges	-6.64	(0.02)	***	-6.85	(0.02)	***	-9.54	(0.12)	***	-9.50	(0.13)	***
Homophily												
Skin color (nodematch)	0.62	(0.02)	***	0.25	(0.03)	***	0.15	(0.03)	***	0.38	(0.07)	***
Gender (nodematch)							1.08	(0.03)	***	1.08	(0.03)	***
Grade (nodematch)							2.54	(0.03)	***	2.54	(0.03)	***
Parents' completed education (nodematch)							0.14	(0.03)	***	0.14	(0.03)	***
Race (nodematch, diff)												
Black				1.00	(0.04)	***	1.59	(0.07)	***	1.47	(0.08)	***
Latino				0.80	(0.04)	***	0.63	(0.07)	***	0.65	(0.07)	***
Asian				1.83	(0.16)	***	1.29	(0.19)	***	1.39	(0.20)	***
Native American				1.87	(0.16)	***	1.39	(0.20)	***	1.33	(0.21)	***
White				0.94	(0.03)	***	0.69	(0.06)	***	0.69	(0.06)	***
Other				0.55	(0.35)		-0.21	(0.37)		-0.20	(0.37)	
Missing				–	–	–	–	–	–	–	–	–
Nodal Attributes												
Skin color							-0.03	(0.02)	†	-0.05	(0.02)	**
Race (ref: Black)												
Latino							0.39	(0.05)	***	0.42	(0.05)	***
Asian							0.40	(0.06)	***	0.51	(0.07)	***
Native American							0.37	(0.07)	***	0.34	(0.08)	***
White							0.32	(0.05)	***	0.32	(0.05)	***
Other							0.67	(0.06)	***	0.68	(0.07)	***
Missing							-0.43	(0.25)	†	-0.39	(0.32)	
Edge Attributes												
Activity Comembership							2.38	(0.03)	***	2.38	(0.03)	***
Network Structure												
Mutuality							–	–	–			
Transitivity							–	–	–			
Open triads												
Indegree (gw)							–	–	–			
Outdegree (gw)							–	–	–			
Skin Color Homophily x Ego Race												
Black (ref)												
Latino										-0.16	(0.05)	***
Asian										-0.41	(0.10)	***
Native American										-0.02	(0.10)	
White										-0.12	(0.05)	***
Other										-0.14	(0.11)	
Missing										-0.20	(0.50)	
BIC	**99695**			**98577**			**82728**			**82795**		

Source: Teen Identity Development and Education Study. Notes: * p<0.05; ** p<0.01; *** p<0.001.

Table 10. Southwest Results of the Exponential Random Graph Models of Friendship at Wave 2

	Model 1			Model 2			Model 3			Model 4		
Effects	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***
Edges	-6.91	(0.02)	***	-7.09	(0.02)	***	-9.49	(0.13)	***	-9.43	(0.13)	***
Homophily												
Skin color (nodematch)	0.61	(0.03)	***	0.25	(0.03)	***	0.13	(0.03)	***	0.41	(0.08)	***
Gender (nodematch)							0.92	(0.03)	***	0.92	(0.03)	***
Grade (nodematch)							2.66	(0.03)	***	2.66	(0.03)	***
Parents' completed education (nodematch)							0.19	(0.03)	***	0.20	(0.03)	***
Race (nodematch, diff)												
Black				0.85	(0.04)	***	1.32	(0.07)	***	1.17	(0.08)	***
Latino				0.84	(0.04)	***	0.59	(0.07)	***	0.59	(0.07)	***
Asian				1.55	(0.20)	***	0.59	(0.23)	***	0.69	(0.23)	***
Native American				2.23	(0.12)	***	1.82	(0.15)	***	1.78	(0.16)	***
White				0.91	(0.04)	***	0.79	(0.06)	***	0.81	(0.07)	***
Other				0.27	(0.33)		-0.16	(0.35)		-0.16	(0.35)	
Missing				–	–	–	–	–	–	–	–	–
Nodal Attributes												
Skin color							-0.08	(0.02)	†	-0.10	(0.02)	***
Race (ref: Black)												
Latino							0.37	(0.05)	***	0.39	(0.05)	***
Asian							0.43	(0.06)	***	0.56	(0.07)	***
Native American							0.33	(0.06)	***	0.32	(0.07)	***
White							0.20	(0.05)	***	0.20	(0.05)	***
Other							0.45	(0.06)	***	0.44	(0.07)	***
Missing							-0.39	(0.28)		-0.25	(0.34)	
Edge Attributes												
Activity Comembership							2.47	(0.03)	***	2.47	(0.03)	***
Network Structure												
Mutuality							–	–	–			
Transitivity							–	–	–			
Open triads												
Indegree (gw)							–	–	–			
Outdegree (gw)							–	–	–			
Skin Color Homophily x Ego Race												
Black (ref)												
Latino										-0.17	(0.05)	***
Asian										-0.48	(0.10)	***
Native American										-0.10	(0.09)	
White										-0.17	(0.05)	***
Other										-0.08	(0.11)	
Missing										-0.53	(0.60)	
BIC	**96749**			**95784**			**80125**			**80188**		

Source: Teen Identity Development and Education Study. Notes: * p<0.05; ** p<0.01; *** p<0.001.

Table 11. Southwest Results of the Exponential Random Graph Models of Friendship at Wave 3

	Model 1			Model 2			Model 3			Model 4		
Effects	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***	**Estimate**	***(SE)***	***p***
Edges	-7.01	(0.02)	***	-7.15	(0.02)	***	-10.05	(0.15)	***	-10.00	(0.16)	***
Homophily												
Skin color (nodematch)	0.63	(0.03)	***	0.34	(0.03)	***	0.20	(0.03)	***	0.52	(0.10)	***
Gender (nodematch)							0.88	(0.03)	***	0.88	(0.03)	***
Grade (nodematch)							2.60	(0.04)	***	2.60	(0.04)	***
Parents' completed education (nodematch)							0.13	(0.03)	***	0.13	(0.03)	***
Race (nodematch, diff)												
Black				0.65	(0.05)	***	1.26	(0.09)	***	1.07	(0.11)	***
Latino				0.64	(0.05)	***	0.45	(0.08)	***	0.46	(0.08)	***
Asian				1.82	(0.18)	***	0.71	(0.21)	***	0.80	(0.21)	***
Native American				2.15	(0.12)	***	1.78	(0.17)	***	1.72	(0.19)	***
White				0.78	(0.04)	***	0.61	(0.07)	***	0.61	(0.08)	***
Other				0.02	(0.45)		-0.59	(0.46)		-0.58	(0.46)	
Missing				–	–	–	–	–	–	–	–	–
Nodal Attributes												
Skin color							0.01	(0.02)		-0.01	(0.02)	
Race (ref: Black)												
Latino							0.49	(0.06)	***	0.52	(0.06)	***
Asian							0.64	(0.07)	***	0.75	(0.08)	***
Native American							0.41	(0.07)	***	0.39	(0.08)	***
White							0.43	(0.06)	***	0.44	(0.06)	***
Other							0.64	(0.07)	***	0.64	(0.09)	***
Missing							0.48	(0.22)	*	0.69	(0.26)	***
Edge Attributes												
Activity Comembership							2.44	(0.04)	***	2.43	(0.04)	***
Network Structure												
Mutuality							–	–	–			
Transitivity							–	–	–			
Open triads												
Indegree (gw)							–	–	–			
Outdegree (gw)							–	–	–			
Skin Color Homophily x Ego Race												
Black (ref)												
Latino										-0.18	(0.06)	***
Asian										-0.40	(0.11)	***
Native American										-0.07	(0.11)	
White										-0.18	(0.06)	***
Other										-0.14	(0.13)	
Missing										-0.78	(0.48)	
BIC		**70313**			**69880**			**58968**			**59039**	

Source: Teen Identity Development and Education Study. Notes: * p<0.05; ** p<0.01; *** p<0.001.

Table 12. Confidence intervals of interaction effects between skin color homophily and race

	Confidence Intervals	
	Lower Bound	Upper Bound
Midwest		
Skin Color Homophily x Ego Race		
Wave 1		
Latino	0.02752636	0.4854548
Asian	0.03119108	0.3246952
Native American	0.22552560	1.1013517
White	0.04357676	0.3248388
Other	0.11010055	0.6262080
Wave 2		
Latino	-0.26313844	0.1047934
Asian	0.06057877	0.3113540
Native American	0.11517280	0.9533262
White	0.04813256	0.2911240
Other	0.31287204	0.7378200
Wave 3		
Latino	-0.01029666	0.4152910
Asian	0.28084402	0.6060368
Native American	-0.18498510	0.8809785
White	0.26924010	0.5822333
Other	0.22640968	0.7466854
Southwest		
Skin Color Homophily x Ego Race		
Wave 1		
Latino	0.13640904	0.3140960
Asian	-0.21308493	0.1675248
Native American	0.15950480	0.5598826
White	0.17668320	0.3394319
Other	0.04523645	0.4448198
Wave 2		
Latino	0.14946605	0.3318395
Asian	-0.26467164	0.1367638
Native American	0.14243830	0.4903244
White	0.15398675	0.3333569
Other	0.13240435	0.5387784
Wave 3		
Latino	0.21620995	0.4496460
Asian	-0.09398999	0.3231091
Native American	0.23823430	0.6507042
White	0.22594735	0.4549884
Other	0.13704950	0.6105524

Source: Teen Identity Development and Education Study.

Paper 3: Skin Color Stratification in Educational Contexts: An Intersectional Approach

Abstract:

The formal, systematic study of racial inequality in schools has a long-standing history in the social sciences reaching back to the mid-twentieth century. Prior research, however, is limited in its ability to assess how racial appearance plays a unique role in structuring unequal outcomes among students. Using skin color data captured through an innovative data collection design, the present study aims to address this gap in the literature through an intersectional lens by examining to what extent does student skin color influence experiences of discrimination from school authority figures and peers, levels of academic achievement, and academic efficacy. Results show that skin color is statistically significant predictor of school-based and peer-based discrimination and academic achievement, however it does not impact academic efficacy. Further while significant associations do not differ by gender, they are conditioned on racial self-identification. To be sure, findings reveal that White and Asian students are particularly less susceptible to skin color bias in schools, relative to their Black peers, while Latino and Native American students and students identifying as being of some "other" race experience similar skin color penalties.

Introduction

In 1954 the U.S. Supreme Court ruled that separate educational facilities for black and white students were inherently unequal, thereby requiring that segregated schools be phased out throughout the nation. While the 1954 ruling strongly asserted the unconstitutional nature of the "separate but equal" doctrine put forth by segregationists, it did not clearly declare the rate at which educational integration should proceed, but rather vaguely described that states should comply with "all deliberate speed." This nebulose timeline ultimately resulted in a tremendous backlash that made desegregation a slow and arduous process prompting the NAACP, the leading civil rights organization that brought the *Brown v. Board of Education* case to the Supreme Court, to organize a group of students to desegregate Little Rock Central High School. This group of students, who came to be known as the Little Rock Nine, had to be escorted by armed military to step foot on school grounds due to the verbal and physical attacks of which they were the target.

While the U.S. school system is no longer riddled with examples of such unbridled, overt racism and discrimination like that which was experienced by the Little Rock Nine and their contemporaries, social scientists have since been asserting that covert forms of prejudice and discrimination in schools persists as evidenced by unequal outcomes. Further, an under investigated dimension of race upon which racism and prejudicial attitudes operate is that of racial appearance, which research shows has the potential to play a pivotal role in determining individual life chances. The under investigation of this often-overlooked dimension of race therefore has the potential to mask the extent of racial inequity in schools presently.

The present study thus aims to go beyond the common lines of inquiry undertaken by education scholars by examining the effects of student skin color on school-based and peer-based discrimination, academic achievement, and academic efficacy–four outcomes of educational import–and uses an intersectional lens to test for whether such associations are moderated by ethnoracial identification and gender.

Background

Skin Color Stratification

Colorism or phenotypic bias are concepts that are used to describe prejudice and discrimination based on racial phenotypic markers such as skin, hair and eye color, hair texture, and facial features such as eye, nose, and lip shape. Colorism as a process operates so that features considered to be prototypically Eurocentric, such as light skin, hair and eye color, smooth hair, serve as a form of social capital thereby providing those who hold such features social advantage over those who do not. Indeed, a large body of work has documented the effects of colorism by providing evidence of skin color stratification across a host of sociological outcomes. Examples of such research date as far back as Myrdal's seminal work on race relations in the U.S., wherein he found that Black individuals of light skin tones obtained a higher occupational status than those of darker skin tones (1944). Since then, similar findings corroborating the advantages of lighter skin-tone have become increasingly common place throughout the literature. Specifically, differences in skin color have been linked to unequal outcomes including those pertaining to education, income, wealth, housing, the labor market, and the criminal justice system (Arce, Murguia, and Frisbie 1987; Murguia and Telles 1996; Espino and Franz, 2002; Hill 2000; Bodenhorn 2006; Gyimah-Brempong and Price 2006; Kizer 2017).

Skin Color Stratification and Education

Education is not immune to colorism and the subsequent unequal outcomes the system of oppression produces. Several scholars have undertaken the task of showing how individuals of lighter skin tone fair better than their darker-skinned counterparts in different ways within educational settings. Spearheading this effort, Hughes and Hertel used data from the National Survey of Black Americans (NSBA) to examine, among other things, the life chances of Black Americans ultimately finding that the gap in years of education between whites and blacks was close to identical to the gap between light- and dark-skinned blacks. Verna M. Keith and Cedric Herring (1991) extended Hughes and Hertel's work by more accurately quantifying the differences between skin-tone groups using NSBA data, ultimately showing that a one-unit increase in skin-tone gradient rating was associated with an additional half-year increase in completed education. Comparisons on the polar ends of the skin-tone gradient used in the NSBA showed that those on the light end (i.e., very light-skin-tone) were found to have attained more than two additional years of education than those attributed with having "very dark brown complexion" (Keith and Herring 1991). Murguia and Telles find similar results among Mexican-Americans with similar family backgrounds (1996). Simply put, Mexican-Americans of darker complexion were found to have lower levels of educational attainment than those with lighter skin.

Research examining skin tone stratification in education extends beyond examining skin color differences in educational attainment. Studies evincing skin color disparities in schools include those examining differences in school discipline, grade point average (GPA), teacher expectations, and peer interactions (Hunter 2007). For example, darker-skinned students of color are more likely to experience more severe school discipline, and considering increasing police presence in schools, subsequently experience a greater push toward the school-to-prison pipeline than students of color with lighter complexion (Hannon, DeFina, & Bruch 2013). An analysis of National Longitudinal Study of Adolescent Health (Add Health) and Adolescent Health and

Academic Achievement (AHAA) data show that individuals of darker skin tone receive significantly lower GPA than lighter-skinned peers–a finding robust to investigation at both the within- and between-racial group level (Thompson and McDonald 2016).

While quantitative data on teacher expectations and skin color is scant, a larger body of literature on how race conditions teacher expectations, measured as racial identification, exists (Masten et al. 1999; Wynne 1999). For example, in their study on teacher expectations, Masten and colleagues (1999) asked elementary school teachers to rate White and Hispanic 5th grade students in the areas of learning, motivation, creativity, and leadership. Findings showed that Anglo-American (White) students were rated higher across these attributes than Hispanic students, and highly acculturated Hispanic students received higher ratings than their Hispanic peers deemed as having lower levels of acculturation. A few qualitative studies help to support these findings (Fergus 2009; Marx 2002). Fergus, in particular, discovered a link between racial appearance and teacher expectations in his interviews with 17 Mexican and Puerto Rican high school students. Specifically, study subjects revealed how they believed the way they racially appeared phenotypically influenced the way they were identified by teachers, and subsequently, played a role in their treatment by teachers (Fergus 2009). If teachers and other school authorities perceive students of lighter skin tones to be more intelligent, from better families, and more well-behaved than their darker-skinned counterparts, students may adjust to meet those racialized expectations (Murguia and Telles 1996).

Expectations from school authority figures are not the only expectations that have the potential to influence educational outcomes among students. Indeed, an entire body of interdisciplinary research is focused on the effects peer influence has on adolescent lives (e.g., Bauman and Ennett 1996; Maxwell 2002; Steinberg and Monahan 2007; Brown et al. 2008;

Schaefer, Haas and Bishop 2012; Van Zalk et al. 2010; Sherman et al. 2016; Wildman et al. 2016). Further, that school peers have been found to express some of the same prejudicial attitudes held by teachers coupled with what is already known about how peer perceptions influence adolescent behavior suggests that potential discrimination from peers based on racial appearance may also negatively impact adolescents. To be sure, it has been argued that students of color often ascribe their lighter-skinned peers as being more attractive, intelligent, and having greater social status, relative to their darker-skinned peers (Craig 2002; Robinson and Ward 1995; Torres 2006). Thus, following the aforementioned line of reasoning, it is important to *empirically* test for whether skin color is a determinate of peer-based discriminating.

Variations by Ethnoracial Group

While the vast majority of work investigating racial discrimination ignores the potential presence of colorism, the literature on colorism has typically investigated the phenomenon within ethnoracial groups individually with far less examining skin color inequality across ethnoracial groups directly. For instance, Bailey, Saperstein and Penner interrogate differences in income inequality by racial self-classification, skin color, and a combination of both measures across nine-teen countries in the Americas ultimately finding that while in some countries income inequality could be understood best by race, or by skin color alone, some country's income inequality, like the United States', could be best understood by a combination of both (2014). Several studies with similar designs, albeit concerned with different inequality outcomes, can be found in the literature (Monk 2016; Kelly 2020; Bailey, Loveman, and Muniz 2013; Perreira and Telles 2014). While these studies have undoubtedly improved our understanding of the value added by considering both racial self-classification and skin color as predictors of different types of inequality, few have assessed whether differences in the impact of skin color on inequality exist by ethnoracial groups (Hunter 2016; Thompson and McDonald 2016).

Variations by Gender

The idea that colorism or skin color bias operates differently for women and men is not new. Skin tone is an important physical marker that plays a salient role in defining what is considered beautiful and beauty is an important form of capital for women (Hunter 2007; Hunter 2002; Bond and Cash 1992; Dixon and Telles 2017). Beauty provides women with social status that can benefit them within their occupations and education, and even the allow them to fair better in the marriage market (Hunter 2005; Hamilton, Goldsmith and Darity 2009). Whereas skin color is closely tied to cultural beliefs about beauty and attractiveness among women, it is more closely tied to ideas about intelligence and criminality among men since men are viewed as the gender that relies more on their laurels than their looks (Farley, Chia and Allred 1998; Freedman 1986; Lakoff and Scherr 1984; Gyimah-Brempong and Price 2006). Although it is commonly understood that gender dynamics are at play in conditioning the association between skin color and various meaningful outcomes of sociological significance, few studies investigate whether these differences are statistically significant.

Aiming to address these gaps in the literature, this paper explores the main and intersecting effects of self-identified race/ethnicity and perceived skin color on levels of perceived school-based and peer-based discrimination, academic grades, and levels of academic efficacy among students between- and within-self-identified Whites, Blacks, Latinos, Native Americans, and Asians and whether the association between skin color and the outcomes of interest differ between adolescent boys and girls.

Data and Methods

To answer these quandaries the present study draws on data from two unique components of the Teen Identity Development and Education Study (TIDES)–a longitudinal study that's original focus was to examine the role that peer relationships play in adolescents' ethnoracial identity and in turn, how ethnoracial identity and peer relationships influence educational outcomes among students. The first component of the study involved survey data collection from 9^{th}-12^{th} grade students located at two ethnically diverse high schools. Data was collected at both high schools at three different points in time over a one-year period: Spring 2017, Fall 2017, and Spring 2018. Each of the high schools sampled are geographically located in two distinct U.S. regions–in the Southwest (SW) and Midwest (MW)– resulting in a combined overall sample of N_{W1} = 3,191, N_{W2} = 3,605, N_{W3} = 3,109. While both schools are characterized as being ethnically and racially diverse, they are diverse in unique ways. To be sure, the TIDES sample is 21% (MW) and 29% (SW) Black, 7% (MW) and 27% (SW) Latino/Hispanic, 24% (MW) and 3% (SW) Asian, 1% (MW) and 4% (SW) Native American, 41% (MW) and 33% (SW) White, and 05% (MW) and 4% (SW) identified as being of some other race. Gender was evenly distributed with a minor overrepresentation among boys in the SW school (SW=52% and MW=46%) and a slightly minor overrepresentation among girls in the MW school (SW=47% and MW=51%). Both schools had 2% of subjects who identified as being of some "other" gender.

Skin Color Coding Procedure

All measures included in the present analyses were collected during school hours via paper survey, except for skin color. Self-reported skin color data was not collected via paper survey in an effort to reduce measurement error, as prior research has found that perceptions of skin color may be skewed depending on the ethnoracial identity of the individual doing the perceiving (Hill 2002). Therefore, rather than asking respondents to rate their own skin color, an innovative data collection design was created as the second component of the study wherein subjects' skin color was coded with the use of high school yearbook photos. Yearbooks were obtained from both schools at the end of each academic year and were digitized. The vast majority of subject images within yearbooks were headshot photos, however, the SW 2017-2018 yearbook included full body images of graduating seniors. Student headshots were individually cropped and placed into a computer slide show so that each headshot or body image photo appeared on its own individual slide with an associated image identification number. Once this process was complete, the slideshow was randomized, and an all-black slide was placed between each slide containing a subject image. Slide shows were then broken up into smaller slide shows containing approximately 80-100 images each to help regulate the time raters spent coding. Data entry was completed by coders directly using the Qualtrics survey platform. Coders were instructed to open Qualtrics window and to size the window so that they could see both the survey and the digitized image queue containing subjects' photos on the screen at the same time. Through the Qualtrics survey platform coders were presented with three skin color guides/palettes: the NIS Skin Color Scale also known as the "Scale of Skin Color Darkness" (Massey and Martin, 2003), the PERLA Color Palette (Telles, 2008), and the SCAP Color Scale (Bond and Cash, 1992; adapted by Gonzales-Backen and Umaña-Taylor, 2011). Each guide/palette was presented separately, and coders were instructed to select the color that most resembled the skin color of the individual in the image they were presently coding. A total of 38 coders coded photos over the course of four academic quarters. Coders were all undergraduate students from the University of California, Irvine, except for one coder who was a graduate student. Every headshot was coded by no less than three coders.

Table 13 presents the cross-tabulation between skin tone and race. Not surprisingly, a one-way ANOVA revealed significant clustering across both variables (F-value=1,409; p<0.001). This is particularly true among White subjects—most of whom are clustered on the lighter end of the spectrum—whereas there is considerably more variability among the other racial groups.

(TABLE 13 ABOUT HERE)

Measures

In what follows, we describe the measures used in the analysis and their respective coding.

<u>School-based discrimination</u>: Using an adapted version of the Adolescent Discrimination Distress Index (ADDI; Fisher, Wallace, and Fenton 2000) a scale was developed to measure the perceived experiences of racial/ethnic discrimination from school authorities. Items included in the scale include responses to the following questions: 1) Were you put in a lower ability class or group because of your race/ethnicity?; 2) Were you disciplined unfairly or given school detention because of your race/ethnicity?; and 3) Were you given a lower grade than you deserved because of your race/ethnicity? Students were permitted to respond to these questions using a Likert scale where 1=Never, 2=Once or twice, 3=A few times, 4=A lot, and 5=A whole lot. The scale produced a Cronbach's alpha ranging from 0.68-0.84.

<u>Peer-based discrimination</u>: Also using an adapted version of the ADDI, a four-item scale was developed to assess the degree to which students felt they were being discriminated against by peers. Items in the scale include responses to the following questions: 1) Were you called insulting names by other kids because of your race/ethnicity?; 2) Were you threatened by other kids because of your race/ethnicity?; 3) Did other kids exclude you from their activities because

of your race/ethnicity?; and 4) Did other kids assume your English was poor because of your race/ethnicity? The scale produced a Cronbach's alpha ranging from 0.72-0.75.

Academic grades: To assess academic achievement, students self-reported their academic grades by answering the question, "Thinking about your grades across all of your subjects **THIS YEAR**, what are your grades generally in school?" Nine response categories were provided to students from which they could choose ranging from 1=Mostly A's: 100%-90%-9=Mostly F's: 59% or below. Self-reported grades were only collected in Waves 2 and 3, resulting in larger rates of missingness when data were combined across waves for analyses.

Academic efficacy: Academic efficacy was also measured using an adapted scale that draws from items in the Patterns of Adaptive Learning Scales (PALS; Midgley et al. 2000). PALS scale items have been developed and refined over time by a group of researchers using goal orientation theory to examine the relation between the learning environment and students' motivation, affect, and behavior. Items in the scale used in the current study include how much students agree with the following statements: 1) I'm certain I can master the skills taught in class this year; 2) I'm certain I can figure out how to do the most difficult class work; 3) I can do almost all the work in class if I don't give up; 4) Even if the work is hard, I can learn it; and 5) I can do even the hardest work in this class if I try. Students were instructed to think about how true they believed these statements to be within the past 30 days. A five-point Likert scale was used for the response categories provided to survey participants. Items on the student scales are anchored at 1 =Not at all true, 3 =Somewhat true, and 5 =Very true. Assessments on the internal consistency of the scale used in this study produced alphas ranging from 0.91-0.93.

Skin Color: Although several skin color measures exist TIDES dataset, the PERLA skin color measure was chosen because it has been suggested that it may have advantages over the

Massey-Martin scale in studies that examine individuals with lighter skin tones (Gordon 2022). The SCAP scale measure was not considered as a viable skin color measure for this analysis because it does not represent darker skin tones commonly found among black populations. The PERLA color palette is an 11-item chart intended to capture variations in skin color found in Latin America with an emphasis on the darker end of the color spectrum. Lower values in the 11-item chart represent lighter tones and higher values represent darker tones. To help diminish potential bias, skin color was operationalized by averaging all PERLA color ratings.

Racial self-identification: Although there are three waves of data, we assume race to be temporally invariant and use the self-identified race at first mention to racially categorize individuals. Racial categories include “Black/African American”, “Latino or Hispanic”, “Asian”, “American Indian/Native American”, “White”, and an “Other” category.

Gender: Student gender is measured using self-report and is also assumed to be temporally invariant and so first mentioned gender is used in the current analyses. Three gender categories are included in the current study: 1=male, 2=female, and 3=some other gender.

Analytic Strategy

Linear regression models are used to assess whether skin color is associated with perceived school-based and peer-based discrimination, academic grades, and academic efficacy. Models were specified using a stepwise process. Each of the specified models also control for survey wave, yearbook from which subject headshot was coded for skin color, grade level, highest level of parental education, and immigrant generation. Models were also specified to produce standardized coefficients for all continuous variables.

Results

Associations Between Skin Color and School-Based Discrimination

Table 14 presents estimates of associations between subject skin tone and perceived discrimination at the hands of school authority figures. Model 1 represent a baseline model that only includes skin color as a key predictor and also includes controls for survey wave, yearbook, grade, highest level of parent education, and immigrant generation. Model 2 includes skin color, self-identified race and gender as predictors and include controls for survey wave, yearbook, grade, highest level of parent education, and immigrant generation. Models 3 adds a moderating effect between skin color and self-identified race and Model 4 adds a moderating effect between skin color and gender. Results reveal statistically significant positive associations between skin tone and school-based perceived discrimination across all four models at the $p<0.001$-level. Model 2, the model wherein a control for racial self-identification is introduced, shows that Asian, Native American, and Non-Hispanic White students, all report lower levels of school-based discrimination on average, than Non-Hispanic Black students, however Latino/Hispanic students and those who identified as being of some "other" race did not differ significantly from their Black counterparts. Gender is also introduced in Model 2 revealing students who identify being of some "other" gender are more likely to report perceiving discrimination from school authority figures, than their male counterparts. Being female, however, does not predict perceiving experiences of school-based discrimination.

Model 3 of Table 14 tests for an interacting effect between skin color and ethnoracial group revealing a statistically significant moderating effect between skin color the White ethnoracial category at the $p<0.05$-level. That the moderating effect is negative suggests that the strength of the association between skin color and the likelihood of perceiving school-based discrimination is weaker for White students, relative to Black students. In other words, in terms of predicting levels of perceived school-based discrimination, the effect of skin color is uniform for students across all ethnoracial groups, except for White students.

Model 4 of Table 14 specifies an interacting effect between skin color and student gender showing no statistically significant moderating effect between skin color and gender at the $p<0.05$-level; however, a marginal positive effect between our key predictor and being of some “other” gender is evident. Overall, Table 2 shows that skin color has a robust effect on perceptions of school-based discrimination, beyond that of self-identified race and that this association differs for White students but not for students of any particular gender.

(TABLE 14 ABOUT HERE)

Associations Between Skin Color and Peer-Based Discrimination

Table 15 presents estimates of the associations between adolescent skin tone and peer-based discrimination. As seen in Table 14, results reveal statistically significant positive associations between student skin tone and peer-based discrimination across all models at at least the $p<0.05$ level. Model 1 shows a statistically significant coefficient for skin tone of 0.17; however, once a control for self-identified race is added to our models, as seen in Model 2, we see a non-trivial decrease to our standardized coefficient at 0.07. Self-identified race is again a statistically significant predictor of perceptions of peer-based discrimination. To be sure, Model 2 shows, Latinos/Hispanics, Native Americans, and Non-Hispanic Whites, were less likely to report perceiving discrimination from peers, relative to Non-Hispanic Black students. Gender is also introduced in Model 2 revealing much like perceiving discrimination from school authority figures students who identify being of some “other” gender are more likely to report perceiving discrimination from peers than male students.

Model 3 includes our moderating effect between skin color and self-identified race, again revealing a significant effect between skin tone and the Asian and White ethnoracial racial category. That the moderating effect is negative shows that the strength of the association between skin color and peer-based discrimination is weaker for Asian and White students relative to Black students. Model 4 tests for interacting effects between skin color and gender showing no significant moderating effects between both predictors suggesting that gender does not condition skin color bias at the student level.

(TABLE 15 ABOUT HERE)

Associations Between Skin Color and Self-Reported Grades

Table 16 shows the estimates of associations between student skin tone and academic achievement, measured as self-reported grades. All models presented in Table 16, on average, show statistically significant negative associations between skin color and adolescents' self-reported grades. Model 4 is the only model in which only marginal significance is achieved. That all other models show statistically significant associations between skin tone and self-reported grades at the $p<0.05$-level strongly suggests that among students the darker their skin tone, the lower their academic achievement. Model 2 reveals that skin color results are generally robust to the inclusion of self-identified race and gender as key predictors. To be sure, being Asian, being White, or being of some "other" race is associated with having higher grades, relative to the Black reference category. While identifying with either the female gender or some "other" gender is associated with self-reported grades the direction of the association differs between groups. Whereas being a girl is associated with having higher grades, relative to boys, being of some "other" gender is associated with having lower grades relative to boys. Models 3 and 4 reveal no statistically significant interaction effects between skin color and self-identified race and gender, respectively, aside from a marginally significant positive interaction effect between skin color and the Non-Hispanic White group. Overall Table 16 reveals that skin color plays a unique role in predicting academic achievement among students beyond that of self-identified race and gender.

(TABLE 16 ABOUT HERE)

Associations Between Skin Color and Academic Efficacy

Table 17 shows results for models predicting levels of academic efficacy among students. Unlike results shown in Tables 14-16, skin color is not a statistically significant predictor of academic efficacy any of the four specified models. Both race and gender, however, are evidently robust predictors across all analyses in which they are included (i.e., Models 2-4). To be sure, both being Asian or being White is associated with having higher levels of academic efficacy. Further, results show that girls have an association with higher levels of academic efficacy than boys, while students identifying as being of some “other” gender is associated with lower levels of academic efficacy, relative to boys. Results shown in Table 17 are the only models throughout the set of analyses in the current study that do not reveal a link between skin color and an educational outcome of interest.

(TABLE 17 ABOUT HERE)

Discussion

This study compares educational outcomes among students differentially located in intersecting, multidimensional status hierarchies on the axes of skin tone, self-identified race and gender to better understand the way in which these hierarchies operate to produce inequality in the school environment. Overall, findings from the regression analyses provide evidence suggesting skin color bias is mechanism driving unequal outcomes among students, above and beyond bias based on racial self-identification, gender, and an array of other socio-demographic controls. That skin color was associated with school- and peer-based discrimination and academic achievement but not associated with academic efficacy suggests that bias on the basis of racial phenotypic appearance is directly disadvantaging darker-skinned students through unfair treatment and assessment, as opposed to impacting belief in their academic capabilities subsequently worsening their performance. Prior research's inclination to examine singular educational outcomes per empirical study and that those outcomes typically measure academic performance or achievement without considering other important educational outcomes has limited understanding of the mechanisms shaping skin color disparities between students.

Using reliable skin color data collected by way of an innovative data collection design which leveraged high school yearbook photos–we analyzed associations between student skin color and several meaningful educational outcomes of interest. Results showed that skin color is significantly associated with experiences of discrimination at the hands of school authority figures *and* peers and with academic grades and these associations were not being driven by self-identified race. Findings are in line with research on the multidimensionality of race and add further evidence in support of race scholars' calls to consider differing aspects of race, particularly those of phenotypic appearance, in investigations of racial inequity (Roth 2016).

We observed evidence that the association between skin color and perceptions of school- and peer-based discrimination was conditioned on race, however race did not differentiate the association between skin color and academic grades. White students particularly experience a weaker association between their skin tone and their experiences of school-based discrimination, while both Asian and White students show weaker associations between skin color and discrimination from peers–relative to Black students. Findings align with prior research on stereotype promise, showing that Asian-Americans, many of darker skin tones and of South-East Asian descent, procure advantage from being associated with high-achieving, light-skinned Asians of East-Asian descent (Lee and Zhou 2015) and a large body of literature on how whiteness serves as form of racial capital to protect individuals who are perceived as authentically claiming the identity from disadvantage (Doane 2003; Roediger 2005; Lipsitz 2006). In other words, while having darker skin appears to penalize students of all ethnoracial groups, on average, the penalty is either reduced or eliminated entirely for White and Asian students, revealing an understudied advantage among these higher-achieving groups.

No intersecting effects on the axes of gender were evident in the association between skin color and any of the educational outcomes under investigation. Although students fitting outside the gender binary were more likely to report experiencing discrimination from school officials and peers, were more likely to report having poorer grades, and were more likely to report lower levels of academic efficacy than both boys and girls, having darker skin tone does not place them at a greater disadvantage. Girls appear to be at a disadvantage relative to boys with respect to their grades, however they fair better in terms of their belief in their efficacy levels regarding academic abilities–with girls' skin tone not playing a major role in either association. Ultimately findings reveal that while gender dynamics may have a hand in how colorism partially constructs beauty standards (e.g., Hunter 2002; Dixon and Telles 2017) or who society views as criminal (Gyimah-Brempong and Price 2006), it does not directly influence how colorism operates in school environments.

That individuals of darker skin color, on average, perceive higher levels of discrimination from teachers and other school officials, higher levels of discrimination from their peers, the majority of which they encounter in school, are more likely to earn poorer grades, but fair similarly with regard to their levels of academic efficacy relative to their lighter-skinned counterparts suggests a certain degree of resilience in the face of significant adversity. Intervention efforts aiming to mitigate such inequality should therefore not center changing the behavior of dark-skinned students, but rather emphasize the color bias that exists and ways in which that bias may be reduced through diversity training among teachers, other school personnel and students alike.

Conclusion

An important distinction between studies on colorism and racial discrimination is that the former concerns itself with skin tone and other racial markers of physical appearance while the latter concerns itself with racial identification regardless of physical appearance. Few studies, however, examine how the two inextricably linked and often overlapping forms of oppression covertly operate to produce unequal outcomes among students and even fewer studies, if any, examine if these intersecting hierarchies operate differently for boys, girls and non-binary students. Findings from the current study reveal that, although challenging, future studies aspiring to unveil the inner workings of ethnoracial hierarchies are hard-pressed to do so through an intersectional lens.

Paper 3: References

Bailey, Stanley R., Mara Loveman, and Jeronimo O. Muniz. 2013. “Measures of ‘Race’ and the Analysis of Racial Inequality in Brazil.” *Social Science Research* 42(1):106–19.

Bauman, Karl E., and Susan T. Ennett. 1996. “On the Importance of Peer Influence for Adolescent Drug Use: Commonly Neglected Considerations.” *Addiction* 91(2):185–98.

Bond, Selena, and Thomas F. Cash. 1992. “Black Beauty: Skin Color and Body Images among African-American College Women.” *Journal of Applied Social Psychology* 22(11):874–88.

Brown, B. Bradford, Jeremy P. Bakken, Suzanne W. Ameringer, and Shelly D. Mahon. 2008. “A Comprehensive Conceptualization of the Peer Influence Process in Adolescence.” *Understanding Peer Influence in Children and Adolescents* 13:17–44.

Craig, Maxine Leeds. 2002. *Ain’t I a Beauty Queen? Black Women, Beauty, and the Politics of Race*. Oxford ; New York: Oxford University Press.

Dixon, Angela R., and Edward E. Telles. 2017. “Skin Color and Colorism: Global Research, Concepts, and Measurement.” *Annual Review of Sociology* 43(1):405–24.

Doane, Woody. 2013. “Rethinking Whiteness Studies.” Pp. 11–26 in *White out*, edited by A. W. Doane and E. Bonilla-Silva. Routledge.

Farley, Sally D., Rosina C. Chia and Linda J. Allred. 1998. "Stereotypes about Attractiveness: When Beautiful is Not Better." *Journal of Social Behavior and Personality* 13(3):479-492.

Fisher, Celia B., Scyatta A. Wallace, and Rose E. Fenton. 2000. “Discrimination Distress during Adolescence.” *Journal of Youth and Adolescence* 29(6):679–95.

Freedman, Rita 1986. Beauty Bound. Lexington, MA: Lexington Books.

Gonzales-Backen, M. A., & Umana-Taylor, A. J. (2011). Examining the role of physical appearance in Latino adolescents’ ethnic identity. Journal of Adolescence, 34(1), 151–162.

Gyimah-Brempong, Kwabena, and Gregory N. Price. 2006. "Crime and Punishment: And Skin Hue Too?" *American Economic Review* 96(2):246–50.

Hamilton, Darrick, Arthur H. Goldsmith, and William Darity Jr. 2009. "Shedding 'Light' on Marriage: The Influence of Skin Shade on Marriage for Black Females." *Journal of Economic Behavior & Organization* 72(1):30–50.

Hannon, Lance, Robert DeFina, and Sarah Bruch. 2013. "The Relationship Between Skin Tone and School Suspension for African Americans." *Race and Social Problems* 5(4):281–95.

Hill, Mark E. 2002. "Race of the Interviewer and Perception of Skin Color: Evidence from the Multi-City Study of Urban Inequality." *American Sociological Review* 67(1):99–108.

Hughes, Michael, and Bradley R. Hertel. 1990. "The Significance of Color Remains: A Study of Life Chances, Mate Selection, and Ethnic Consciousness among Black Americans." *Social Forces* 68(4):1105–20.

Hunter, Margaret L. 2002. "'If You're Light You're Alright': Light Skin Color as Social Capital for Women of Color." *Gender and Society* 16(2):175–93.

Hunter, Margaret L. 2005. *Race, Gender, and the Politics of Skin Tone*. New York: Routledge.

Hunter, Margaret. 2007. "The Persistent Problem of Colorism: Skin Tone, Status, and Inequality." *Sociology Compass* 1(1):237–54.

Hunter, Margaret. 2016. "Colorism in the Classroom: How Skin Tone Stratifies African American and Latina/o Students." *Theory Into Practice* 55(1):54–61.

Keith, Verna M., and Cedric Herring. 1991. "Skin Tone and Stratification in the Black Community." *American Journal of Sociology* 97(3):760–78.

Kelly, Monique DA. 2020. "Examining Race in Jamaica: How Racial Category and Skin Color Structure Social Inequality." *Race and Social Problems* 12(4):300–312.

Lakoff, Robin and Racquel Scherr 1984. Face Value: The Politics of Beauty. Boston, MA: Routledge.

Lee, Jennifer, and Min Zhou. 2015. *The Asian American Achievement Paradox*. New York: Russell Sage Foundation.

Lipsitz, George. 2006. *The Possessive Investment in Whiteness: How White People Profit from Identity Politics*. Temple University Press.

Marx, Sherry. 2002. "Entanglements of Altruism, Whiteness, and Deficit Thinking." Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Massey, Douglas S., and Jennifer A. Martin. 2003. The NIS Skin Color Scale.

Masten, William G., Maximino Plata, Karen Wenglar, and Joe Thedford. 1999. "Acculturation and Teacher Ratings of Hispanic and Anglo-American Students." *Roeper Review* 22(1):64–65.

Maxwell, Kimberly A. 2002. "Friends: The Role of Peer Influence Across Adolescent Risk Behaviors." *Journal of Youth and Adolescence* 31(4):267–77.

Midgley, Carol, Martin L. Maehr, Ludmila Z. Hruda, Eric Anderman, Lynley Anderman, Kimberley E. Freeman, and T. Urdan. 2000. "Manual for the Patterns of Adaptive Learning Scales." *Ann Arbor: University of Michigan*.

Monk, Ellis P., Jr. 2016. "The Consequences of 'Race and Color' in Brazil." *Social Problems* 63(3):413–30.

Myrdal, Gunnar, Richard Sterner, and Arnold M. Rose. 1944. An American dilemma; the Negro problem and modern democracy. New York: Harper & Brothers.

Perreira, Krista M., and Edward E. Telles. 2014. “The Color of Health: Skin Color, Ethnoracial Classification, and Discrimination in the Health of Latin Americans.” *Social Science & Medicine* 116:241–50.

Perreira, Krista M., Joshua Wassink, and Kathleen Mullan Harris. 2019. “Beyond Race/Ethnicity: Skin Color, Gender, and the Health of Young Adults in the United States.” *Population Research and Policy Review* 38(2):271–99.

Robinson, Tracy L. and Ward, Janie V. 1995. “African American Adolescents and Skin Color.” *Journal of Black Psychology* 21(3):256–74.

Roediger, David R. 2005. *Working toward Whiteness: How Americas Immigrants Became White: The Strange Journey from Ellis Island to the Suburbs*. New York: Basic Books.

Roth, Wendy D. 2016. “The Multiple Dimensions of Race.” *Ethnic and Racial Studies* 39(8):1310–38.

Schaefer, David R., Steven A. Haas, and Nicholas J. Bishop. 2012. “A Dynamic Model of US Adolescents’ Smoking and Friendship Networks.” *American Journal of Public Health* 102(6):12–18.

Sherman, Lauren E., Ashley A. Payton, Leanna M. Hernandez, Patricia M. Greenfield, and Mirella Dapretto. 2016. “The Power of the like in Adolescence: Effects of Peer Influence on Neural and Behavioral Responses to Social Media.” *Psychological Science* 27(7):1027–35.

Telles EE. Project on Ethnicity and Race in Latin America (PERLA) 2013 Retrieved February 11, 2017 from https://perla.princeton.edu/.

Thompson, Maxine S., and Steve McDonald. 2016. “Race, Skin Tone, and Educational Achievement.” *Sociological Perspectives* 59(1):91–111.

Torres, Kimberly C. 2006. “Manufacturing Blackness: Skin Color Necessary But Not Sufficient. Race Relations and Racial Identity at an Ivy League University (Pennsylvania).” Dissertation Abstracts International, University of Pennsylvania.

Van Zalk, Maarten Herman Walter, Margaret Kerr, Susan J. T. Branje, Håkan Stattin, and Wim H. J. Meeus. 2010. “It Takes Three: Selection, Influence, and de-Selection Processes of Depression in Adolescent Friendship Networks.” *Developmental Psychology* 46(4):927–38.

Widman, Laura, Sophia Choukas-Bradley, Sarah W. Helms, and Mitchell J. Prinstein. 2016. “Adolescent Susceptibility to Peer Influence in Sexual Situations.” *Journal of Adolescent Health* 58(3):323–29.

Wynne, Joan T. 1999. “The Elephant in the Living Room: Racism in School Reform.” Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Table 13. Cross-tabulation, skin color by race

	Race						
Skin Color	Black	Latino/ Hispanic	Asian	Native American	White	Other	Total
1							
n	0	16	1	2	103	1	123
Percent	0.0	1.8	0.2	1.5	6.0	0.5	2.7
2							
n	4	60	12	4	512	7	599
Percent	0.4	6.7	2.3	3.0	30.1	3.6	13.1
3							
n	7	143	41	11	575	32	809
Percent	0.6	16.0	7.8	8.2	33.8	16.7	17.7
4							
n	41	373	180	59	374	45	1072
Percent	3.6	41.7	34.4	44.0	22.0	23.4	23.4
5							
n	131	222	167	41	114	55	730
Percent	11.5	24.8	31.9	30.6	6.7	28.6	15.9
6							
n	287	67	73	13	20	34	494
Percent	25.3	7.5	13.9	9.7	1.2	17.7	10.8
7							
n	251	8	27	0	4	7	297
Percent	22.1	0.9	5.2	0.0	0.2	3.6	6.5
8							
n	220	3	18	3	1	9	254
Percent	19.4	0.3	3.4	2.2	0.1	4.7	5.5
9							
n	119	2	5	1	0	0	127
Percent	10.5	0.2	1.0	0.7	0.0	0.0	2.8
10							
n	56	0	0	0	0	1	57
Percent	4.9	0.0	0.0	0.0	0.0	0.5	1.2
11							
n	20	0	0	0	0	1	21
Percent	1.8	0.0	0.0	0.0	0.0	0.5	0.5
Total							
N	1136	894	524	134	1703	192	4,583
Percent	100	100	100	100	100	100	100

Source: Teen Identity Development and Education Study

Table 14. Linear regression models predicting frequency of school-based perceived discrimination by skin color (PERLA)

	Model 1: Skin Color			Model 2: Skin Color + Race + Gender			Model 3: Skin Color * Race			Model 4: Skin Color * Gender		
	Estimate	***p***	**(SE)**	**Estimate**	***p***	**(SE)**	**Estimate**	***p***	**(SE)**	**Estimate**	***p***	**(SE)**
Skin color	0.24	***	(0.02)	0.14	***	(0.03)	0.20	***	(0.04)	0.14	***	(0.03)
Race (ref: Non-Hispanic Black)												
Latino/Hispanic				-0.13	†	(0.04)	-0.06		(0.08)	-0.13	†	(0.07)
Asian				-0.31	***	(0.04)	-0.25	**	(0.09)	-0.31	***	(0.08)
Native American				-0.33	**	(0.05)	-0.26	*	(0.11)	-0.34	**	(0.10)
Non-Hispanic White				-0.34	***	(0.04)	-0.34	***	(0.08)	-0.35	***	(0.07)
Other				-0.27	†	(0.08)	-0.20		(0.15)	-0.27	†	(0.15)
Gender (ref: Male)												
Female				0.03		(0.03)	0.02		(0.03)	0.03		(0.03)
Other				0.53	***	(0.12)	0.53	***	(0.12)	0.60	***	(0.12)
Skin color X Race												
Latino/Hispanic							-0.06		(0.08)	–		–
Asian							-0.05		(0.08)	–		–
Native American							0.01		(0.16)	–		–
Non-Hispanic White							-0.14	*	(0.07)	–		–
Other							0.11		(0.22)	–		–
Skin color X Gender												
Female							–		–	-0.01		(0.03)
Other							–		–	0.22	†	(0.13)
N	3,542		–	3,489		–	3,489		–	3,489		–
Multiple R-squared	0.07			0.08			0.08			0.08		
Adjusted R-squared	0.06		–	0.07		–	0.07		–	0.07		–

Source: Teen Identity Development and Education Study

Notes: † p<0.10; * p<0.05; ** p<0.01; *** p<0.001. All models control for survey wave, yearbook from which subject headshot was coded for skin color, gender, grade, highest level of parent education, and immigrant generation. All continuous predictors report standardized coefficients.

Table 15. Linear regression models predicting frequency of peer-based perceived discrimination by skin color (PERLA)

	Model 1: Skin Color			Model 2: Skin Color + Race + Gender			Model 3: Skin Color * Race			Model 4: Skin Color * Gender		
	Estimate	***p***	**(SE)**	**Estimate**	***p***	**(SE)**	**Estimate**	***p***	**(SE)**	**Estimate**	***p***	**(SE)**
Skin color	0.17	***	(0.02)	0.07	*	(0.03)	0.17	***	(0.03)	0.08	*	(0.03)
Race (ref: Non-Hispanic Black)												
Latino/Hispanic				-0.27	***	(0.07)	-0.16	*	(0.08)	-0.27	***	(0.07)
Asian				-0.05		(0.08)	0.09		(0.09)	-0.04		(0.08)
Native American				-0.46	***	(0.11)	-0.34	**	(0.11)	-0.46	***	(0.11)
Non-Hispanic White				-0.33	***	(0.07)	-0.26	**	(0.08)	-0.33	***	(0.07)
Other				-0.12		(0.15)	-0.01		(0.16)	-0.12		(0.15)
Gender (ref: Male)												
Female				0.04		(0.03)	0.05		(0.03)	0.04		(0.03)
Other				0.31	**	(0.12)	0.32	**	(0.12)	0.25	*	(0.13)
Skin color X Race												
Latino/Hispanic							-0.09		(0.08)	–		–
Asian							-0.25	**	(0.08)	–		–
Native American							0.13		(0.16)	–		–
Non-Hispanic White							-0.16	*	(0.07)	–		–
Other							-0.23		(0.22)	–		–
Skin color X Gender												
Female							–		–	-0.02		(0.03)
Other							–		–	-0.21		(0.13)
N	3,541		–	3,488		–	3,488		–	3,488		–
Multiple R-squared	0.06			0.07			0.07			0.07		
Adjusted R-squared	0.05		–	0.06		–	0.07		–	0.06		–

Source: Teen Identity Development and Education Study

Notes: † p<0.10; * p<0.05; ** p<0.01; *** p<0.001. All models control for survey wave, yearbook from which subject headshot was coded for skin color, gender, grade, highest level of parent education, and immigrant generation. All continuous predictors report standardized coefficients.

Table 16. Linear regression models predicting grades by skin color (PERLA)

	Model 1: Skin Color			Model 2: Skin Color + Race + Gender			Model 3: Skin Color * Race			Model 4: Skin Color * Gender		
	Estimate	***p***	**(SE)**	**Estimate**	***p***	**(SE)**	**Estimate**	***p***	**(SE)**	**Estimate**	***p***	**(SE)**
Skin color	-0.16	***	(0.02)	-0.06	*	(0.03)	-0.14	*	(0.07)	-0.10	†	(0.05)
Race (ref: Non-Hispanic Black)												
Latino/Hispanic				0.03		(0.07)	-0.02		(0.12)	0.04		(0.11)
Asian				0.59	***	(0.07)	0.90	***	(0.14)	0.96	***	(0.12)
Native American				0.00		(0.10)	-0.06		(0.17)	0.00		(0.16)
Non-Hispanic White				0.36	***	(0.07)	0.63	***	(0.13)	0.58	***	(0.11)
Other				0.25	**	(0.09)	0.38	*	(0.16)	0.39	**	(0.15)
Gender (ref: Male)												
Female				-0.17	***	(0.03)	-0.27	***	(0.05)	-0.27	***	(0.05)
Other				-0.38	***	(0.11)	-0.61	***	(0.18)	-0.67	***	(0.19)
Skin color X Race												
Latino/Hispanic							-0.04		(0.12)	–		–
Asian							0.14		(0.13)	–		–
Native American							-0.04		(0.26)	–		–
Non-Hispanic White							0.17	†	(0.10)	–		–
Other							-0.22		(0.18)	–		–
Skin color X Gender												
Female							–		–	0.02		(0.05)
Other							–		–	-0.15		(0.19)
N	3,234		–	3,234		–	3,234		–	3,234		–
Multiple R-squared	0.15			0.18			0.18			0.18		
Adjusted R-squared	0.15		–	0.17		–	0.18		–	0.17		–

Source: Teen Identity Development and Education Study

Notes: † p<0.10; * p<0.05; ** p<0.01; *** p<0.001. All models control for survey wave, yearbook from which subject headshot was coded for skin color, gender, grade, highest level of parent education, and immigrant generation. All continuous predictors report standardized coefficients.

Table 17. Linear regression models predicting academic efficacy by skin color (PERLA)

	Model 1: Skin Color			**Model 2: Skin Color + Race + Gender**			**Model 3: Skin Color * Race**			**Model 4: Skin Color * Gender**		
	Estimate	***p***	**(SE)**	**Estimate**	***p***	**(SE)**	**Estimate**	***p***	**(SE)**	**Estimate**	***p***	**(SE)**
Skin color	-0.02		(0.02)	0.05	†	(0.03)	0.05		(0.04)	0.05		(0.03)
Race (ref: Non-Hispanic Black)												
Latino/Hispanic				0.11	†	(0.06)	0.10		(0.07)	0.11	†	(0.07)
Asian				0.27	***	(0.07)	0.24	**	(0.08)	0.27	***	(0.07)
Native American				0.15		(0.10)	0.15		(0.10)	0.15		(0.10)
Non-Hispanic White				0.21	**	(0.07)	0.20	**	(0.07)	0.21	**	(0.07)
Other				0.17	†	(0.09)	0.17	†	(0.09)	0.17	†	(0.09)
Gender (ref: Male)												
Female				0.09	**	(0.03)	0.08	**	(0.03)	0.09	**	(0.03)
Other				-0.38	**	(0.12)	-0.34	**	(0.11)	-0.35	**	(0.13)
Skin color X Race												
Latino/Hispanic							-0.03		(0.07)	–		–
Asian							0.04		(0.08)	–		–
Native American							0.14		(0.15)	–		–
Non-Hispanic White							0.00		(0.06)	–		–
Other							-0.10		(0.11)	–		–
Skin color X Gender												
Female							–		–	-0.01		(0.03)
Other							–		–	0.07		(0.14)
N	4,045		–	4,045		–	4,045		–	4,045		–
Multiple R-squared	0.03			0.03			0.03			0.03		
Adjusted R-squared	0.02		–	0.03		–	0.03		–	0.03		–

Source: Teen Identity Development and Education Study

Notes: † p<0.10; * p<0.05; ** p<0.01; *** p<0.001. All models control for survey wave, yearbook from which subject headshot was coded for skin color, gender, grade, highest level of parent education, and immigrant generation. All continuous predictors report standardized coefficients.

CONCLUSION

Summation of Dissertation Findings

This dissertation aimed to uncover how a failure to account for racial appearance, measured as skin color, in studies examining racial stratification within school contexts has the potential to mask inequality among students. The dissertation not only aimed to address how accounting for skin color impacted student educational outcomes but also problematized the issue of how best to capture skin tone data.

To be sure, the first paper of the dissertation examined the reliability and construct validity of some of the most widely used skin color instruments across the social sciences. Using an extended component of the TIDES project, which entailed coding of adolescents' yearbook headshots for skin color by multiple raters, this paper asked: (1) How do three varying, widely-used skin color scales measure up in terms of their reliability and construct validity when modeling outcomes involving adolescent perceived experiences of discrimination; (2) Does one of the skin color measures provide greater unique explanatory power beyond that of racial self-identification above others when predicting different forms of discrimination; and (3) Does the predictive validity of each of the skin color scales differ for different ethnoracial groups? Results showed no substantial evidence that one scale should be preferred over the other included in the analyses in terms of their respective reliability and construct validity. Additionally, results indicated that each of the skin color scales have predictive power beyond that of racial self-classification across all forms of perceived discrimination with very few statistically significant differences between them. Differences in predictive validity between scales among Black, Latino, Asian, and White subjects were evident. Overall scales revealed the greatest predictive validity when they align best with the range of the skin tones a given ethnoracial population holds.

The second paper of the dissertation investigated whether skin color serves as an axis of racial differentiation that contributes to the extent of racial segregation in adolescent friendship networks. Using sociocentric friendship network data from two diverse high schools in two distinct geographic regions and Exponential Random Graph Models this paper asked: 1) Are adolescents similar in skin color more likely to be tied to one another within friendships, net of homophily based on racial self-classification and 2) Is there evidence that the extent of skin color homophily differs across ethnoracial groups? Findings showed evidence of skin color homophily above and beyond that of racial homophily and significant differences in the extent of skin color homophily by ethnoracial group. The study's findings underscore the importance of operationalizing race in multiple dimensions when assessing racial exclusion in social networks to avoid yielding conservative estimates.

The third and final paper of the dissertation used an intersectional lens to interrogate several ways in which skin color stratifies educational outcomes among students. Simply put, Paper 3–which also used TIDES data–questioned the extent to which student skin color was associated with experiences of discrimination from school authority figures and peers, levels of academic achievement and academic efficacy and also whether associations differed by ethnoracial group or gender. Results showed that skin color is statistically significant predictor of school-based and peer-based discrimination and academic achievement, however it did not influence levels of perceived academic efficacy among students. Further while significant associations did not differ by gender, they were conditioned on racial self-identification. To be sure, findings revealed that White and Asian students were particularly less susceptible to skin color bias in schools, relative to their Black peers, while Latino and Native American students and students identifying as being of some “other” race experience similar skin color penalties.

Future of Research Agenda

The future direction of this research agenda aims to expand upon findings from Paper 2 to assess whether skin color homophily in adolescent friendship networks exacerbates the academic achievement gap among students. To elaborate, while a swath of studies examining educational inequality consider the impact adolescents have on their peers in academic settings, less do so using structural, relational approaches. Social network analysis as a methodological approach allows for investigation of social structure, through the use of graph and probability theory, by giving precise formal definitions to aspects of the social structural environment (Wasserman and Faust 1994). Further, advances in network modeling and an increase in longitudinal relational data collection provide advantages in answering the ever-present question of endogeneity; namely whether adolescents are influencing one another overtime thereby becoming more alike, or if adolescents who are similar to begin with select into friendships with one another. The ability to disaggregate the effect of peer influence from homophilous selection, in addition to understanding other components of network evolution, is made possible by statistical modeling of longitudinal network data collected in panel design, where repeated observations of a constant set of network actors are available.

Looking ahead, this research agenda will use these models, known as Stochastic Actor-Oriented Models (SAOMs) or Sienna models to answer the following questions: 1) What is the relative contribution of racial homophily and skin color homophily (net of racial homophily) in adolescent friendship networks on academic achievement, measured as student self-reported grades? Drawing on literature on racial stratification and cumulative disadvantage, the proposed study helps push forth the field's understanding of how racial and color assortativity within adolescent friendships contribute to academic disparities within schools by way of peer influence.

Broader Impacts

This research agenda endeavors to make theoretically and methodologically laden contributions to the field of sociology, particularly within the quantitatively oriented sub-discipline of social network analysis. Additionally, because social network analysis is an area in which racial and ethnic minorities and women are severely underrepresented, investigation of the social aspects of race and gender within the field are limited. Pursuing network research through an intersectional lens will therefore provide a fuller account of the mechanisms driving social inequality and will ultimately help to identify key intervention strategies that will benefit our multicultural, multigendered society.