CrossMark

Promotive and Corrosive Factors in African American Students' Math Beliefs and Achievement

Matthew A. Diemer[1,2,3] · Aixa D. Marchand[2] · Sarah E. McKellar[2] ·
Oksana Malanchuk[3]

Abstract Framed by expectancy-value theory (which posits that beliefs about and the subjective valuation of a domain predict achievement and decision-making in that domain), this study examined the relationships among teacher differential treatment and relevant math instruction on African American students' self-concept of math ability, math task value, and math achievement. These questions were examined by applying structural equation modeling to 618 African American youth (45.6 % female) followed from 7th to 11th grade in the Maryland Adolescent Development in Context Study. While controlling for gender and prior math achievement, relevant math instruction promoted and teacher differential treatment corroded students' math beliefs and achievement over time. Further, teacher discrimination undermined students' perceptions of their teachers, a mediating process under-examined in previous inquiry. These findings suggest policy and practice levers to narrow opportunity gaps, as well as foster math achievement and science, technology, engineering and math success.

Introduction

Students' beliefs about their math ability and the importance they place upon math both play key roles in science, technology, engineering and mathematics (STEM) achievement and occupational decision-making (e.g., Adelman 1999; Wang and Degol 2013). Expectancy-value theory frames student motivation and academic success as largely predicted by academic self-concept (i.e., domain-specific beliefs about ability) and task value (i.e., the subjective importance of a task) (e.g., Eccles and Wigfield 2002). As students navigate academic contexts, changes in self-concept of math ability, math task value, and math achievement have been linked to teachers' practices and social interactions (Urdan and Schoenfelder 2006). Student–teacher interactions can play a promotive or corrosive role in motivating students and enhancing achievement (Eccles and Roeser 2011; Watt 2006). This study examines the role of "relevant math instruction," or math teachers' emphases upon making lessons interesting and the usefulness of math for students' everyday life. While research has established the relationships between these math teaching practices and student math-related beliefs (e.g., Watt 2004), there is less known about whether relevant math instruction fosters these same outcomes among African American students (Eccles 2005; Okeke et al. 2009).

Furthermore, for students of color, the school racial climate—conceptualized here as students' perceptions that

✉ Matthew A. Diemer
    diemerm@umich.edu

    Aixa D. Marchand
    admarch@umich.edu

    Sarah E. McKellar
    smkella@umich.edu

    Oksana Malanchuk
    oksana@umich.edu

1   Room 4120, School of Education, University of Michigan,
    Ann Arbor, MI 48109, USA

2   Combined Program in Education and Psychology,
    Educational Studies, University of Michigan, Ann Arbor, MI,
    USA

3   Research Center for Group Dynamics, Institute for Social
    Research, University of Michigan, Ann Arbor, MI, USA

students of different racial/ethnic groups are treated equally by school personnel (e.g., Benner and Graham 2013)—may have particularly strong implications for students' math beliefs and achievement (Eccles 2005; Graham 1994). The extant literature has understandably focused on first establishing linkages between differential treatment by race and general academic beliefs and outcomes among students of color (e.g., Benner et al. 2015; Chavous et al. 2008). This study advances this literature by examining whether African American students' experiences of unequal treatment based on race by their teachers matters for student outcomes in math. Thus, we will focus on math beliefs and achievement (domain-specific), as compared with prior research that examines academic beliefs and achievement more broadly (domain-general).

From an equity perspective, African American students exhibit lower levels of math achievement than other racial/ethnic groups—which is caused by multiple intersecting factors—yet necessitates further understanding of predictors of African American students' math achievement (McGee and Martin 2011). African American students are also underrepresented in STEM fields and encounter stereotypes that undermine their math achievement (Wang and Degol 2013). Understanding how relevant math instruction and school racial climates promote or corrode math beliefs and achievement, therefore, also has implications for issues of equity and justice.

These issues are examined longitudinally (i.e., from seventh to eleventh grade) among a sample of African American students from Prince George's County, Maryland. This county provides a unique context to study school racial climate and African American math achievement because the socioeconomic status of African American and White families in Prince George's County is relatively comparable, which avoids confounding race and social class (Gutman et al., in press). Prince George's County also had relatively high proportions of African American teachers (36 %) and students (67 %), which is a distinct demographic representation from predominantly White or predominantly African American and Latino/a schools (Cook et al. 2002).

Specifically, this study examines the roles of relevant math instruction and school racial climate in self-concept of math ability, math task value, and math achievement among African American students. Before delving into specifics of the model tested, we first detail the theoretical framework and literature related to key constructs in this study.

Theoretical Framework: Expectancy-Value Model

This study is grounded in the expectancy-value framework (Eccles 1994; Eccles and Wigfield 2002). We examine student self-concept of math ability and math task value

beliefs because of their importance of math for academic success (in general) and for STEM success (in particular) (Adelman 1999; Frank et al. 2008; Simpkins et al. 2006). The expectancy-value model also provides conceptual space to integrate relevant math instruction and school racial climates—which correspond to the "socializers' beliefs and behaviors" and "cultural milieu" components of the expectancy-value model—and serve as precursors of self-concept of math ability, task value, and achievement (Eccles 1994). Eccles (2005) emphasized the importance of context and culture in academic beliefs and achievement: "We believe, however, that these [achievement-related] choices are heavily influenced by socialization pressures and cultural norms" (p. 10). Thus, the expectancy-value framework is well-suited to address our research questions (Wigfield et al. 2015). This study examines how student beliefs are related to interactions with their teachers as part of the classroom context.

Student Motivational Beliefs

Self-Concept of Math Ability

Self-concepts of ability, or beliefs about one's competency in a given domain, play key roles in achievement and in educational/occupational decision-making in the expectancy-value framework (Eccles 1994; Eccles and Wigfield 2002; Wang and Degol 2013). Self-concepts of ability, like self-efficacy beliefs, are domain-specific—youth have differentiated self-concepts of ability for math versus reading, for example (Davis-Kean et al. 2008; Sáinz and Eccles 2012). These self-concepts of ability become less inflated by undue optimism over time, as youth experience increased competition and are better able to compare their abilities to others (i.e., they make external comparisons) and to their own abilities in other domains (i.e., they make internal comparisons) (Denissen et al. 2007; Jacobs et al. 2002). Self-concept of math ability is also a strong predictor of math achievement (Denissen et al. 2007; Simpkins and Davis-Kean 2005; Wang and Degol 2013), as evidenced by cross-lagged relationships between self-concept of math ability and math achievement 2 years later (Marsh et al. 2005).

Students' ability beliefs are theoretically distinct from other self-beliefs, such as beliefs about the likelihood of future success in that same domain (i.e., expectancies) as well as beliefs about whether efforts in that domain will produce an intended outcome (i.e., outcome expectations), in social cognitive theories (e.g., Pajares 1996; Eccles and Wigfield 2002). Presumably, these theoretically distinct sets of self-beliefs are influenced by a more global sense of self within that domain (Eccles et al. 1983). Although these concepts are delineated theoretically, empirically self-

concept of ability and expectancies are nearly indistinguishable (Eccles and Wigfield 1995). In this study, we aim to extend prior work by examining factors related to self-concept of math ability among African American students.

Math Task Value

According to the expectancy-value framework, subjective task value encompasses beliefs about how important, interesting, useful, or costly a domain is (Eccles et al. 1983). A key dimension of task value represents the subjective valuation of a particular domain, or how much one values or places importance upon an achievement-related domain, such as mathematics (Gaspard et al. 2015; Watt 2006; Wigfield and Eccles 1992; Wigfield 1994). Task values are domain-specific (e.g., one has a perceived task value for math vs. for sports) and are predictive of choices and achievement in that domain. For example, math task value was predictive of math course enrollment in previous correlational research (Meece et al. 1990; Simpkins and Davis-Kean 2005; Watt 2004, 2006; Wong et al. 2003). Relatedly, field experiments fostering students' science utility value led to subsequent increases in science achievement (Hulleman and Harackiewicz 2009; Hulleman et al. 2010).

Scholars are conceptually and empirically delineating the four specific dimensions believed to underlie subjective task value: intrinsic value (i.e., task is interesting or enjoyable), attainment value (i.e., task aligns with self-schema and/or core personal values), utility value (i.e., task is instrumental for reaching future goals), and cost (i.e., task has anticipated negative consequences) (Canning and Harackiewicz 2015; Durik et al. 2006; Trautwein et al. 2012; Eccles and Wigfield 2002). Yet, the importance or significance of a task is thematic across the first three of these four specific dimensions (i.e., perceived cost is weighed against the subjective importance or significance of intrinsic, attainment, and utility value). While acknowledging the importance of this specificity (an issue returned to in the Limitations section), this study is guided by a more thematic conception of subjective task value, emphasizing importance.

To extend prior work, we seek to identify how contextual factors, namely student–teacher interactions, contribute to perceptions of subjective task value among African American students over time. Self-concepts of ability are theorized to be predictive of subjective task value within the expectancy-value framework, a notion that has received strong empirical support (e.g., Jacobs et al. 2002; Meece et al. 1990). Our study, therefore, examines how contextual factors relate to these relationships over time for African American students.

School and Classroom Contexts: Relevant Math Instruction

In the expectancy-value framework, teachers play a critical role in the development of students' self-concepts of ability and subjective valuation of academic domains (Eccles 1994; Wang and Degol 2013). Relevant math instruction is conceptualized here as the teacher's emphases, throughout the school year, upon making math lessons interesting and relevant to everyday life. The mathematics education and achievement motivation literatures have considered relevant math instruction as a key dimension of the broader authentic math instruction umbrella (Newman and Wehlage 1993; Wang and Eccles 2014). Fostering students' interest in math and emphasizing the relevance of math for everyday life has been associated with students' perceptions of math ability and math task value (Wang and Eccles 2014; Watts 2004; Wigfield and Eccles 2000).

Relevance interventions, which are designed to foster the personal relevance of an academic domain and, thereby, foster motivation and achievement, offer insights into the potential benefits of relevant math instruction: "Making science courses personally relevant and meaningful may engage students in the learning process, enable them to identify with future science careers, foster the development of interest, and promote science-related academic choices (e.g., course enrollment and pursuit of advanced degrees) and career paths" (Hulleman and Harackiewicz 2009, p. 1411). Other field experiments demonstrate that increasing the personal relevance of math is an effective strategy for fostering math task value (Gaspard et al. 2015). While relevance interventions have some key distinctions from teacher instruction, such as duration (i.e., generally "one-shot" interventions in field experiments vs. year-long teacher instructional emphases), this line of inquiry informs our conceptualization of relevant math instruction.

School Racial Climates and General Academic Outcomes

Broadly, school racial climate refers to "school settings' norms and values around race and interracial interactions" (Byrd and Chavous 2011, p. 850). While school racial climate encompasses multiple dimensions, this study focuses on one critical element: how teachers and school personnel interact with students of color. One reason for this focus is that teachers' contributions to school racial climate are more proximal to students' beliefs and achievement than more distal dimensions of school racial climate (i.e., aggregated student perceptions of the overall school climate; Eccles 1994; Eccles and Roeser 2011). This study, therefore, examines the role of teachers in school

racial climates; specifically, the impacts of treating students differently on the basis of race (Chavous 2005; Green et al. 1988).

Relatedly, in the expectancy-value model, teachers are an important determinant of competence beliefs, values, and achievement (e.g., Eccles 1994). Teachers' discrimination severs bonds of students to school and to schooling, undermining the belonging, trust, and connectedness that foster positive academic outcomes and student well-being (Benner et al. 2015; Eccles and Roeser 2011). Thus, teachers' racially discriminatory behaviors may corrode students' academic beliefs, values and achievement (e.g., Eccles and Roeser 2011; Wang and Degol 2013). Negative school racial climates (characterized by higher rates of teacher discrimination) have been associated with lower levels of academic beliefs and performance among students of color (Benner and Graham 2013; Byrd and Chavous 2011; Chavous et al. 2008; Green et al. 1988; Ryan and Patrick 2001). Further, earlier work has observed that more positive school racial climates (in this case, school-level aggregated perceptions of climate) were associated with increases in academic self-efficacy among African American students (Green et al. 1988).

The Current Study

Previous scholarship suggests that relevant instruction promotes, and teacher differential treatment corrodes, domain-general academic beliefs and achievement (e.g., Benner and Graham 2013; Byrd and Chavous 2011; Newmann and Wehlage 1993). This study advances these literatures in two important ways. First, while relevant instruction is associated with academic beliefs and outcomes among predominantly White students (e.g., Newmann and Wehlage 1993; Wang and Eccles 2014), we do not know whether relevant instruction plays the same promotive role among African American students. Secondly, we do not know whether previously identified domain-general effects of teacher differential treatment play a similar role in domain-specific math beliefs, values, and achievement, particularly among African American students. Presumably, corrosive impacts of school racial climate on general academic beliefs and performance also hold true for math-specific beliefs, values, and achievement. According to the expectancy-value model, socializing experiences, cultural contexts, and barriers to opportunity affect the development of self-concepts and task valuation (Eccles 1994; Wang and Degol 2013). Accordingly, positive school racial climates may promote self-concept of math ability and math task value while negative school racial climates may diminish self-concept

of math ability and math task value. These patterns are expected to be more pronounced among African American students, who encounter racialized barriers to the development of competence beliefs and the valuation of math (Simpkins and Davis-Kean 2005; Watt 2004). Further, social influences are particularly pronounced in STEM beliefs and performance (Frank et al. 2008; Wang and Degol 2013), perhaps in part because STEM domains are more stereotyped than other academic domains and/or because of the underrepresentation of people of color in many STEM fields.

This study, therefore, examines the impacts of teacher differential treatment—disciplining students differently on the basis of race, holding lower expectations for students of color, grading students of color more harshly—a key component of school racial climates in school racial climate instrument validation studies (e.g., Green et al. 1988) and recent studies associating school racial climates to academic outcomes (e.g., Benner and Graham 2013; Chavous et al. 2008).

We used structural equation modeling (SEM) to test hypothesized relationships among teacher differential treatment, relevant math instruction, self-concept of math ability, math task value, and math achievement outcomes among African American adolescents followed longitudinally. Math outcomes were measured by adolescents' final math grade in 8th grade, as well as 9th grade Maryland Math standardized achievement scores (detailed further below).

Figure 1 displays the hypothesized paths between the constructs of interest, which are reviewed here. Beginning from the left of the figure (wave three, or 8th grade), teachers' differential treatment is hypothesized to negatively predict relevant math instruction (in that student perceptions of teacher discrimination are hypothesized to undermine student perceptions of relevant math instruction), math grades, self-concept of math ability, math task value, and 9th grade Maryland Math scores. Relevant math instruction is hypothesized to positively predict self-concept of math ability, math task value, 8th grade math grades, and 9th grade Maryland Math scores.

Math task value and self-concept of math ability are hypothesized to have positive autoregressive relationships between 8th and 11th grade. Relevant math instruction and teacher differential treatment are also measured at 8th and 11th grade, yet as these students transitioned from middle to high school and to a new set of teachers, no autoregressive relations were specified between these measures. The same relationships among teacher differential treatment, relevant math instruction, and self-concept of math ability and math task value at 8th grade are also hypothesized in 11th grade.
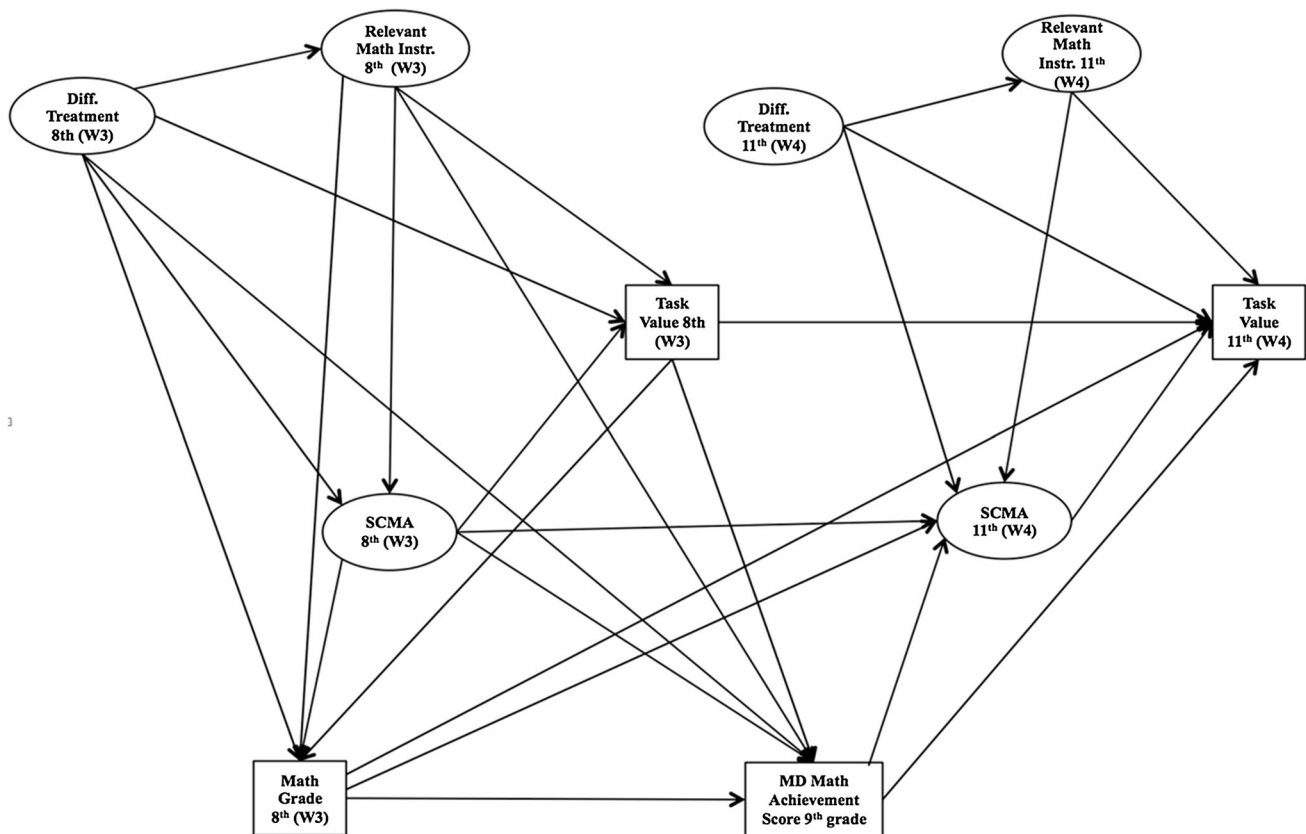
Fig. 1 Hypothesized relations among constructs

Prior academic achievement (i.e., 7th grade math grades) was modeled as a lagged control, and participants' gender was statistically controlled for as well. For clarity, these paths are not depicted in the figure. This lagged measure of math achievement is an ideal control for the later math achievement measures (as well as for other academic constructs) in that it reduces bias in the prediction of the outcome measures and increases confidence in the robustness of inferences drawn (Steiner et al. 2010). While not yielding causal inferences, this strategy (coupled with other approaches that are detailed below) enhances confidence in the inferences drawn from this study.

Methods

Data Source

This study analyzed data from the longitudinal Maryland Adolescent Development in Context Study (MADICS). MADICS sampled children from all 23 public middle schools in Prince George's County, Maryland, a racially and socioeconomically diverse county outside of Washington,

D.C. MADICS is unique in that it sampled African American and White families with comparable levels of socioeconomic status, which affords comparisons across racial groups without confounding social class (Gutman et al., in press). At the time of the MADICS study, 36 % of teachers and 67 % of students in Prince George's County self-identified as African American (Cook et al. 2002).

MADICS began with a sample of 1482 seventh graders at wave 1 in the fall of 1991. This study focuses on data from waves 3 (collected at the end of students' 8th grade in 1993) and 4 (fall of 11th grade in 1996). Waves 3 and 4 were examined because they richly measured youths' perceptions of race, their schools, and their academic abilities.

At Wave 3, the MADICS sample included 1065 adolescents who self-identified as African American ($n = 618$; 58 %), White ($n = 331$; 31 %), or as bi-racial or a member of another minority group ($n = 105$; 10 %). The sample contained more male ($n = 336$; 54.4 %) than female participants ($n = 282$; 45.6 %). Because of our substantive focus on the impacts of school racial climates on math beliefs and achievement, only the African American subsample ($n = 618$) was examined.

General Methodological Approach

SEM simultaneously estimates relations among constructs while adjusting for measurement error and precisely evaluates how well complex models fit the data (Kline 2010). SEM is particularly useful in secondary analyses, where all aspects of a latent construct may not be measured but available indicators can represent that construct. SEM was used to examine how indicators load onto constructs before modeling relations among constructs, to specify and account for measurement error, and to precisely (in that measurement error is better accounted for in latent variable modeling, yielding less biased estimates) test direct and indirect (i.e., mediated) relations (Kline 2010).

The amount of missingness present in these data is detailed in Table 1. All SEM analyses were analyzed under full information maximum likelihood (FIML) conditions, which makes use of all existing data points in analyses, instead of deleting cases listwise or pairwise, an efficient missing data strategy that maximizes effective sample size and statistical power for analyses (Enders 2010). These analyses also included auxiliary variables, which are likely predictors of missing data and attrition (e.g., they measure plausible missingness mechanisms) yet are not directly modeled as predictors or outcome variables. Instead, auxiliary variables "are on the sideline"—they correlate with each other and also predict the residual terms of observed variables in the model (Enders 2010). Auxiliary variables in these analyses included [a] standardized 5th grade California Achievement Test math scores, [b] number of school absences in 7th grade, [c] self-reported likelihood respondent will get involved with drugs later in life, [d] self-reported likelihood respondent will get in trouble with the police later in life, and [e] parental educational attainment (a common measure of household SES; see Diemer et al. 2013). Modeling these as auxiliary variables increases the power of analyses, in that bias due to missingness is attenuated—in addition to accounting for likely predictors of attrition (Muthén and Muthén 2010). It also increases the likelihood of missingness being missing at random (MAR) instead of missing not at random (MNAR), in that likely missingness mechanisms, or sources of attrition (i.e., low achievement, frequent school absences, parental SES, and risky/criminal behavior) are modeled as observed predictors of missingness instead of remaining unobserved sources of bias (Enders 2010).

Measures

Each measure is briefly reviewed below; further detail about each latent construct, observed items, the wording of and response options for each item, and descriptive data is provided in Table 1.

Relevant Math Instruction

Relevant math instruction was measured by student perceptions of how personally meaningful and relevant they perceive their math curriculum to be, consistent with classic (Newmann and Wehlage 1993) and contemporary formulations (Wang and Eccles 2014) of relevant instruction. Student perceptions of instruction are less likely to be influenced by social desirability than teacher perceptions (e.g., teachers might report more favorably on their own teaching practices than student reports of their teacher's practice), and are therefore likely more accurate measures (DeVellis 2003). As depicted in Table 1, the internal consistency of relevant math instruction was acceptable, as measured by Cronbach's alpha. Yet Cronbach's alpha is a misleading estimate of internal consistency (i.e., it is downwardly biased) when a measure consists of few items, because the calculation of alpha is overly sensitive to the number of items in a measure (Clark and Watson 1995; DeVellis 2003). Mean inter-item correlations, which provide more accurate estimates of internal consistency that are not biased by the number of items per measure, are therefore also reported. Generally, good mean inter-item correlation values range from .15 to .50; larger values reflect higher levels of internal consistency (Clark and Watson 1995). The mean inter-item correlations for relevant math instruction suggest high levels of internal consistency (i.e., .52 at wave 3 and .49 at wave 4).

Self-Concept of Math Ability

Self-concept of math ability represents beliefs regarding one's math ability, and was measured by three Likert-type items that are consistent with previous measures of self-concept of academic ability (Denissen et al. 2007; Jacobs et al. 2002; Marsh et al. 2005; Simpkins et al. 2006). As depicted in Table 1, self-concept of math ability was quite internally consistent (as measured by Cronbach alpha and mean inter-item correlations).

Math Task Value

One item measured youths' math task value, which refers to the subjective valuation of math from the expectancy-value framework (Eccles and Wigfield 1995; Watt 2006). Math task value was modeled as an observed indicator in analyses, and as such could not be included in the measurement model (reported below).

Teacher Differential Treatment

Five items measured students' perceptions of differential treatment by teachers (and secondarily, school counselors),

Table 1 Descriptive statistics for variables and observed indicators

Variables/indicators and descriptions	8th grade wave 3					11th grade wave 4				
	M	SD	% Missing	α	IIC	M	SD	% Missing	α	IIC
Self-concept of math ability										
Scale 1–7, example 1 = not at all good to 7 = very good				0.79	0.56				0.82	0.61
v35170/v46038: How good are you in math?	5.07	1.59	1.6			4.64	1.66	22.2		
v35173/v46041: Compared to other kids your age how well do you do in math?	5.1	1.44	1.6			4.86	1.48	29.3		
v35181/v46049: Compared to other kids your age how well do you expect to do next year in math?	5.46	1.32	1.1			5.47	1.33	22.5		
Relevant math instruction										
Scale 1–5, all 1 = almost never to 5 = always				0.68	0.56				0.66	0.49
v35135/v46271: How often does your math teacher use examples that are interesting to you?	3.26	1.22	1.3			3.08	1.27	31.7		
v35137/v46273: How often do you learn things in math that help you with your everyday life?	3.34	1.22	1.6			2.77	1.24	31.6		
Task value (1 item)										
Scale 1–7, all 1 = much less important… 7 = much more important…				N/A	N/A				N/A	N/A
v35177/v46045: Compared to other kids your age, how important are each of the following activities to you: math?	5.29	1.44	1.3			4.96	1.55	29.1		
Perceived differential treatment by race										
Scale 1–5, all 1 = almost never to 5 = always				0.88	0.6				0.88	0.6
v35195/v46327 How often do you feel that teachers call on you less often than they call on other kids because of your race?	1.84	1.12	1.3			1.43	0.86	29.3		
v35196/v46328 How often do you feel that teachers grade you harder than they grade other kids because of your race?	1.69	1.06	1.6			1.4	0.82	29.3		
v35197/v46329 How often do you feel that you get disciplined more harshly be teachers than other kids do because of your race?	1.85	1.17	1.3			1.48	0.87	29.4		
v35198/v46330 How often do you feel that teachers think you are less smart than you really are because of your race?	1.67	1.06	1.3			1.48	0.85	29.3		
v35200/v46331 How often have you felt that teachers/counselors discourage you from taking certain classes because of your race?	1.53	0.96	1.6			1.34	0.79	29.3		
Math grade 7th grade (1 item)										
Scale 1–5	3.37	1.03	4.0							
Math grade 8th grade (1 item)										
Scale 1–5	3.23	0.94	14.0							
Maryland (MD) math achievement—9th grade (1 item)										
Range 273–421	346.69	26.75	21.0							

IIC mean inter-item correlation

on the basis of race. These items were used in previous research to measure students' perceptions of teacher discrimination (e.g., Chavous et al. 2008; Rapa and Diemer 2016; Wong et al. 2003) and are quite similar to the "Equal Status" subscale of the School Interracial Climate Scale (Green et al. 1988), which was informed by intergroup contact theory. The internal consistency of these five items was quite strong, as depicted in Table 1.

Math Achievement

Math achievement during seventh (MADICS wave one) and eighth grade (MADICS wave three) was measured by cumulative math grade for that academic year (on a 1–5 scale). Math grades were obtained from school records. The wave one (7th grade) math achievement measure preceded the wave three (8th grade) and wave four (11th grade) constructs of interest, and therefore wave one math achievement is modeled as a prior control variable in the structural model (please see Fig. 1). Modeling lagged achievement is a particularly strong strategy for addressing unobserved variables bias in educational research, in that prior achievement likely covaries with other observed and unobserved variables (Frank 2000).

Math achievement was also measured by math scores on the Maryland Math Test, a state-level standardized achievement test, which participants completed during the ninth grade (or, between the MADICS waves three [8th grade] and wave four [11th grade]) and was linked to MADICS data. Math achievement during 11th grade was not measured in MADICS, a limitation discussed below.

Results: Measurement Model

CFA first determined the pattern of item loadings onto latent constructs. Items repeated across waves three and four likely share common sources of error variance, and accordingly the residuals of repeated items were estimated and included in all subsequent analyses (Kline 2010). The hypothesized relations between items and their corresponding latent construct—the measurement model—were a very good fit to the data (see Table 2; CFI = .98, TLI = .97, RMSEA = .03).

Further, temporal invariance, or whether latent constructs that were measured at both wave 3 and wave 4 (i.e., relevant math instruction, self-concept of math ability, and differential treatment) mean the same thing and were measured in the same way over time, was examined. (Temporal invariance analyses are similar to establishing measurement invariance across groups, such as gender groups or racial/ethnic groups, but with time essentially serving as the grouping variable.) The first step in this sequential process is establishing configural invariance, or that the "configuration" of observed items loading onto latent constructs was the same at wave 3 and at wave 4. Because analyses indicated that the same items loaded onto the same latent constructs at waves 3 and 4, the criteria for configural invariance were met (Kline 2010; Schmitt and Kuljanin 2008). The next step was establishing loading invariance (also called metric invariance), or that the magnitude of each loading was invariant over time. The loadings for all items comprising self-concept of math

ability and relevant math instruction were invariant across waves 3 and 4, which established "strong invariance" for these measures—that these constructs are measured in the same way, and mean the same thing over time for these participants (Schmitt and Kuljanin 2008).

On the other hand, loadings for teacher differential treatment were **not** invariant over time, suggesting that the measurement of teacher differential treatment changes over time for these participants. This suggests that in the transition from middle school (8th grade in wave 3) to high school (11th grade in wave 4), students' perceptions of racialized mistreatment from teachers changed. It may be that Prince George's County schools maintained an instructional emphasis upon math relevance from middle school to high school—which we cannot examine with these data—but racialized teacher mistreatment changed from middle to high school. This analytic strategy cannot pinpoint whether students transitioned from middle schools with greater levels of differential treatment to high schools with lower levels of differential treatment (suggested by lower differential treatment observed item means at wave 4 than at wave 3, please see Table 1), because the measurement of differential treatment was not equivalent over time. Alternatively, it may be that students' perceptions of what teacher differential treatment means may have changed as these African American students matured. We further consider the potential implications of measurement invariance and non-invariance below.

Structural Model

The structural model tested hypothesized relations among latent constructs, as depicted in Fig. 1, while controlling for gender and for 7th grade (MADICS wave 1) math achievement (not depicted for clarity). Model fit indices indicated that the structural model was a good fit to the data (CFI = .96, TLI = .95, RMSEA = .04). Complete structural model results are reported in Fig. 2 and detailed below (key findings are highlighted in Table 3).

In SEM, standardized coefficients (β) represent effect size estimates, with coefficients of ≥.10 roughly considered "small," .30 to .50 considered "medium," and ≥. 50 considered large (Kline 2010). At wave three (8th grade), teacher differential treatment negatively predicted relevant math instruction (β = −.14) and subsequent 9th grade Maryland Math Achievement scores (β = −.08) and had non-significant relationships to self-concept of math ability, math task value, and math grades in 8th grade. Eighth grade relevant math instruction significantly predicted math task value (β = .20) and self-concept of math ability (β = .48) but a non-significant relationship to 8th grade math grades. Surprisingly, relevant math instruction negatively predicted Maryland Math Achievement scores in

Table 2 Measurement model: factor loadings for latent variables

Latent variable and indicators	Standardized estimate	SE
Self-concept of math ability wave 3		
v35170: How good are you in math?	0.85*	0.03
v35173: Compared to other kids your age how well do you do in math?	0.82*	0.03
v35181: Compared to other kids your age how well do you expect to do next year in math?	0.59*	0.04
Self-concept of math ability wave 4		
v46038: How good are you in math?	0.88*	0.03
v46041: Compared to other kids your age how well do you do in math?	0.89*	0.03
v46049: Compared to other kids your age how well do you expect to do next year in math?	0.57*	0.04
Relevant math instruction wave 3		
v35135: How often does your math teacher use examples that are interesting to you?	0.82*	0.05
v35137: How often do you learn things in math that help you with your everyday life?	0.64*	0.06
Relevant math instruction wave 4		
v46271: How often does your math teacher use examples that are interesting to you?	0.72*	0.05
v46273: How often do you learn things in math that help you with your everyday life?	0.68*	0.06
Differential treatment wave 3		
v35195: How often do you feel that teachers call on you less often than they call on other kids because of your race?	0.69*	0.04
v35196: How often do you feel that teachers grade you harder than they grade other kids because of your race?	0.84*	0.02
v35197: How often do you feel that you get disciplined more harshly be teachers than other kids do because of your race?	0.83*	0.02
v35198: How often do you feel that teachers think you are less smart than you really are because of your race?	0.84*	0.02
V35200: How often have you felt that teachers/counselors discourage you from taking certain classes because of your race?	0.70*	0.04
Differential treatment wave 4		
v46327: How often do you feel that teachers call on you less often than they call on other kids because of your race?	0.69*	0.05
v46328: How often do you feel that teachers grade you harder than they grade other kids because of your race?	0.85*	0.02
v46329: How often do you feel that you get disciplined more harshly be teachers than other kids do because of your race?	0.85*	0.03
v46330: How often do you feel that teachers think you are less smart than you really are because of your race?	0.84*	0.03
v46331: How often have you felt that teachers/counselors discourage you from taking certain classes because of your race?	0.65*	0.06

Significant standardized estimates are denoted with an asterisk

ninth grade ($\beta = -.12$); and was a non-significant predictor of 8th grade math grades. Eighth grade math task value was not a significant predictor of 8th grade math grades or 9th grade Maryland Math scores. Eighth grade self-concept of math ability significantly predicted math grades in eighth grade ($\beta = .29$), 9th grade Maryland Math achievement scores ($\beta = .26$), and 8th grade math task value ($\beta = .54$).

The autoregressive paths between 8th to 11th grade math task value ($\beta = .16$) and self-concept of math ability ($\beta = .34$) at wave three and wave four were significant, yet the magnitude of these estimates suggest some intra-individual change as students transition from middle to high school. [Note: The amount of variance in 11th grade math task value and self-concept of math ability was reduced by the autoregressive relationships specified, and as such one

would expect attenuated estimates of effect size in 11th grade, as compared to 8th grade.] Autoregressive relationships between teacher differential treatment and relevant math instruction were not specified, because participants transitioned from middle school to high school from waves 3–4. Additionally, participants' math grade in 8th grade significantly predicted their score on the Maryland Math test in 9th grade ($\beta = .17$) but was not a significant predictor of math task value or self-concept of math ability in 11th grade.

At 11th grade (wave four), teacher differential treatment was not a significant predictor of relevant math instruction, math task value, or self-concept of math ability. Wave four relevant math instruction significantly predicted self-concept of math ability ($\beta = .47$) and math task value ($\beta = .18$). Ninth grade Maryland Math Achievement
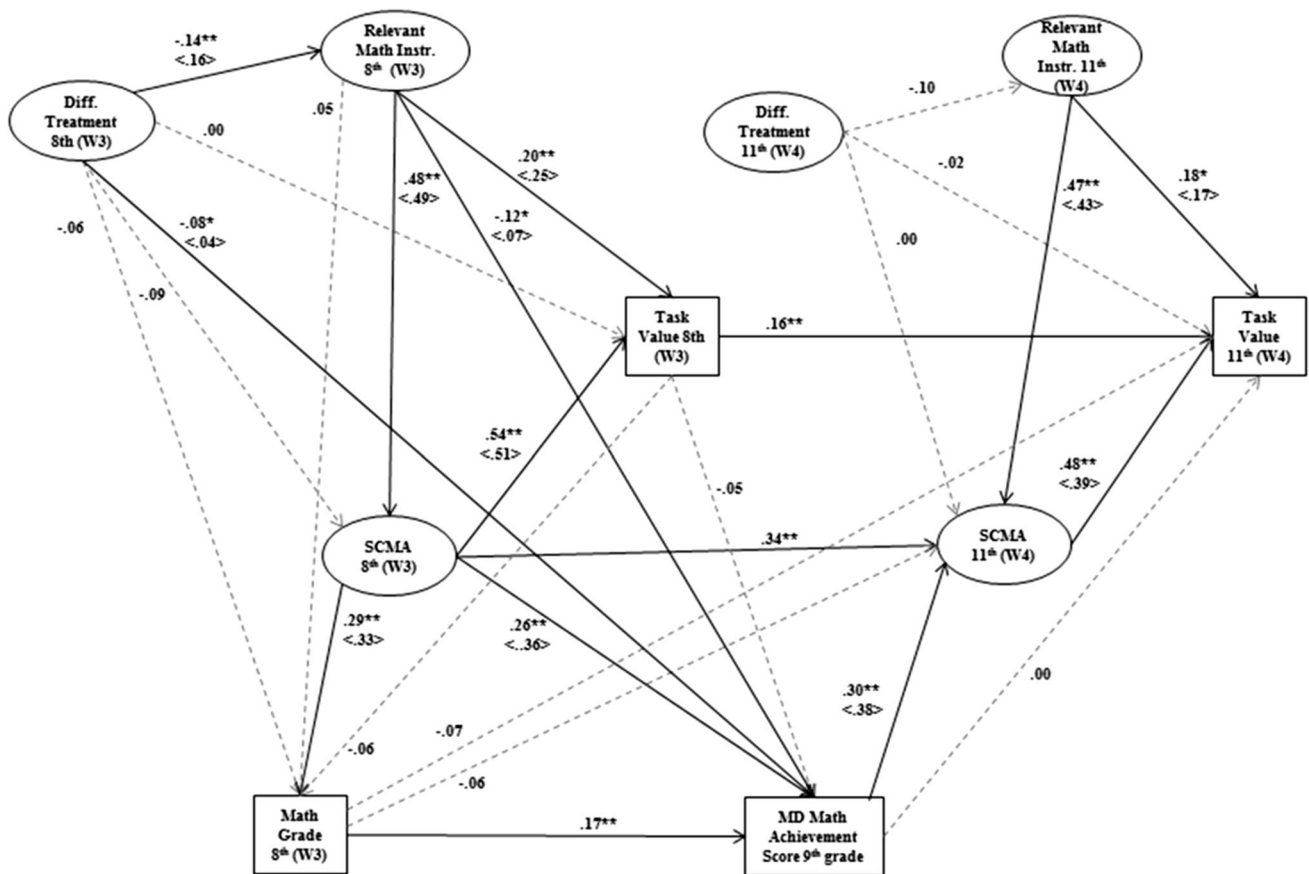
Fig. 2 Structural model. *Note* Paths significant at the .05 level denoted with an *asterisk* and a *solid line*. Non-significant paths are denoted by a *dashed line*. Impact threshold for a confounding variable (ITCV) estimates presented < *inside brackets*>; paths significant at the .01 level are denoted with two *asterisks* and a *solid line*

scores significantly predicted 11th grade self-concept of math ability ($\beta = .30$), but did not significantly predict 11th grade math task value. Eleventh grade self-concept of math ability significantly predicted 11th grade math task value ($\beta = .48$).

A series of substantively-informed mediated relationships were also examined, and are reported in Table 3. Eighth grade self-concept of math ability partially mediated the relationships between 8th grade relevant math instruction and task value at wave three ($\beta = .26$) and between 8th grade relevant instruction and self-concept of math ability at wave four ($\beta = .16$). Similarly, 11th grade self-concept of math ability partially mediated the relationship between 11th grade relevant instruction and math task value ($\beta = .22$).

We also investigated substantively plausible competing models by testing a reverse causality model. (Although a cross-lagged panel model may also be informative, the 3 year time lag between waves 3 and 4 would diminish the capacity of any cross-lagged path to yield robust information about temporal sequencing.) Reverse causality models change the direction of select structural paths to

help us better understand whether construct A "causes" construct B, or vice versa, by comparing the model fit and substantive plausibility of competing models (Pearl 2000). Although reverse causality modeling is helpful in disentangling which construct appears to precede the other, this strategy does not establish causality when applied to observational data. Reverse causality modeling was used here to examine whether self-concept of math ability may instead be predictive of perceptions of relevant math instruction. This entailed reversing the direction of these regressions (relevant math instruction predicts self-concept of math ability in the original structural model depicted in Fig. 1; self-concept of math ability instead predicts relevant math instruction in the reverse causality model) at both wave 3 and at wave 4. To afford comparison, all other paths from the original Fig. 1 are identical in the reverse causality model.

Unexpectedly, this reverse causality model had an equivalent fit with the structural model depicted in Fig. 1 and reviewed above (CFI = .96, TLI = .95, RMSEA = .04). The comparable fit of the reverse causality model indicates that we cannot rule out reciprocal causation between

Table 3 Summary of direct effects and presentation of indirect effects

	Standardized estimate	SE
Hypothesized direct paths		
Differential treatment W3 → relevant math instruction W3	−.14*	.05
Differential treatment W3 → 9th grade math achievement	−.08*	.04
Relevant math instruction W3 → SCMA W3	.48*	.05
Relevant math instruction W3 → task value wave 3	.20*	.06
Relevant math instruction W3 → 9th grade math Achievement	−.12*	.06
SCMA W3 → task value W3	.54*	.06
SCMA W3 → 9th grade math achievement	.26*	.09
SCMA W3 → 8th grade math grade	.29*	.06
Relevant math instruction W4 → task value W4	.18*	.08
Relevant math instruction W4 → SCMA W4	.47*	.06
9th grade math achievement → SCMA W4	.30*	.05
SCMA W4 → task value W4	.48*	.08
Paths from covariates		
Gender → 7th grade math grade	.24*	.04
Gender → differential treatment W3	−.12*	.04
Gender → relevant math instruction W3	−.08	.05
Gender → SCMA W3	−.10*	.04
Gender → task value wave 3	.00	.03
Gender → 8th grade math grade	.12*	.04
Gender → differential treatment W4	−.27*	.05
Gender → relevant math instruction W4	−.11	.06
Gender → SCMA W4	−.06	.05
Gender → task value wave 4	.03	.04
Gender → 9th grade math achievement	−.04	.04
7th grade math grade → differential treatment W3	−.17*	.04
7th grade math grade → relevant math instruction W3	.05	.05
7th grade math grade → SCMA W3	.27*	.04
7th grade math grade → task value wave 3	.06	.04
7th grade math grade → 8th grade math grade	.44*	.04
7th grade math grade → differential treatment W4	−.12*	.06
7th grade math grade → relevant math instruction W4	−.03	.07
7th grade math grade → SCMA W4	.10	.06
7th grade math grade → task value wave 4	.08	.05
7th grade math grade → 9th grade math achievement	.26*	.05
Hypothesized indirect paths		
Relevant math instruction W3 → SCMA W3 → task value W3	.26*	.04
Differential treatment W3 → SCMA W3 → task value W3	−.05	.03
Differential treatment W3 → relevant math instruction W3 → Task Value W3	−.03*	.01
Differential treatment W3 → relevant math instruction W3 → SCMA W3	−.07*	.03
Relevant math instruction W3 → SCMA W3 → SCMA W4	.16*	.03
Differential treatment W3 → relevant math instruction W3 → SCMA W3 → SCMA W4	−.02*	.01
Relevant math instruction W4 → SCMA W4 → Task Value W4	.22*	.05

Significant standardized estimates are denoted with an asterisk

W3 wave 3, *W4* wave 4

teacher's relevant math instruction and students' self-concepts of math ability. It may very well be that students' perceptions of their math ability shape their perceptions of teacher practices; these teacher practices may also foster students' perceptions of math ability. This potentially dynamic process could not be closely examined with these data and indeed, would require micro-data (e.g., multiple close or moment-to-moment observations of both processes) and a wholly different analytic strategy in order to be precisely examined.

In instances where model fit does not provide a clear signal in favor the original or reverse causality model, substantive considerations take on even greater importance. According to expectancy-value theory, teacher practices are specified as prior to students' self-concept of math ability (Eccles 1994, 2005). The relevant instruction literature also specifies teacher instruction as prior to student outcomes (Newmann and Wehlage 1993; Wang and Eccles 2014). Changing the directions of regressions in this reverse causality model also entailed changes to a number of mediating pathways. For example, in this reverse causality model, 8th grade relevant math instruction would now mediate the impact of self-concept of math ability on wave 3 math grades and Maryland Math scores in 9th grade. These changes to mediating pathways, while complex, are less consistent with extant theory than the mediating pathways posited and tested in the Fig. 1 original model. Given the substantive primacy of teacher practices "causing" student outcomes, rather than the reverse, we have therefore decided that the structural model depicted in Fig. 1 and reviewed above is the most substantively plausible, and therefore, final model. However, we return to issues of potential reciprocal causation between relevant math instruction and self-concepts of ability in the Discussion section.

Estimating the Robustness of Inferences

The impact threshold for a confounding variable (ITCV) estimates the magnitude an unobserved variable would need to correlate with a predictor and outcome variable to nullify the statistically significant relationship between that predictor and that outcome (Frank 2000). The ITCV coefficient is a way of estimating how robust an inference is against omitted or unobserved variables bias, yet does not yield causal estimates. In a sense, the ITCV provides a confidence interval around a statistical inference. ITCV values were calculated for substantively important (i.e., not autoregressive relationships) paths and are depicted in brackets in Fig. 2 (in the interest of parsimony, only selected ITCVs will be reviewed here). Note that ITCV values assume that an unobserved confounder is completely uncorrelated with existing covariates, and as such is a conservative estimate (i.e., any unobserved variable

likely shares some variance with existing covariates and as such the "true" value of the ITCV would likely be higher; Frank 2000). In this case, it is unlikely that any unobserved variable has no relationship to the existing covariates—7th grade math achievement and gender—and accordingly these ITCV estimates are likely biased upward.

The obtained ITCV values to nullify the significant relationship between relevant math instruction and self-concept of math ability at 8th grade (.49), as well as at 11th grade (.43), are quite large. By comparison, the ITCV value associated with the path from the 7th grade math grades lagged covariate to 8th grade self-concept of math ability (.27) was substantially smaller. The magnitude of these ITCV coefficients suggests that the paths from relevant math instruction to self-concept of math ability are more robust to unobserved confounding variables (Frank 2000). The ITCV values associated with the significant paths between self-concept of math ability and task value at 8th grade (.51) and at 11th grade (.39) were also quite large, suggesting that these relationships are also more robust to unobserved variables bias. The ITCV values to nullify inferences between 8th grade self-concept of math ability and 8th grade math grades (.33) and between 8th grade self-concept of math ability and 9th grade Maryland Math scores (.36) were fairly large, but not of the same magnitude as the preceding ITCV values. The ITCV value to nullify the inference between 8th grade teacher differential treatment and relevant math instruction (.16) was substantially smaller. This indicates that this inference is less robust—it is more likely that some unobserved variable could nullify this inference than the other significant relationships reviewed above.

Discussion

Framed by expectancy-value theory, this study examined the relationships among teachers' differential treatment and relevant math instruction to African American adolescents' self-concept of math ability, math task value, and math outcomes among a longitudinal sample of youth in more racially and socioeconomically integrated neighborhoods and schools (Gutman et al., in press). These findings support the applicability of expectancy-value theory in framing the development of African American students' math beliefs over time. These findings also support key tenets of the expectancy-value framework, namely, that self-concepts of ability are predictive of task value, and that these beliefs play a role in achievement over time (Denissen et al. 2007; Jacobs et al. 2002; Meece et al. 1990; Watt 2006). Yet, understanding the linkages between these processes among African American students, as well as connecting school racial climates (i.e., teacher differential treatment) to a more

specific academic subject (i.e., math) advances scholarly understanding in two important ways.

The first way is by underscoring the importance of relevant math instruction in the promotion of African American students' math beliefs. While controlling for lagged math achievement and gender and after establishing the temporal invariance of relevant math instruction and self-concept of math ability, relevant math instruction was a significant and strong predictor of self-concept of math ability at 8th and 11th grade—as well as a significant, yet more moderate, predictor of math task value at 8th and 11th grade. In addition to its direct relationship, relevant math instruction also had an indirect (i.e., mediated) relationship to math task value, as self-concept of math ability mediated the relationship between relevant instruction and math task value at 8th and at 11th grade. Establishing the temporal invariance of relevant math instruction and self-concept of math ability also affords more precise comparisons between the magnitude of these relationships at wave three and at wave four, in that these constructs are measured in the same way and mean the same thing for these participants over time (Kline 2010).

Further, 8th grade relevant math instruction was a positive predictor of self-concept of math ability 3 years later, as 8th grade self-concept of math ability mediated the relationship between 8th grade relevant math instruction and 11th grade self-concept of math ability. This suggests that the positive impacts of teachers using math examples that are interesting to students and teaching math in a way that is applicable to everyday life has long-reaching promotive impacts on African American students' math beliefs. These results have some implications for math instruction, particularly for African American students. For example, these findings suggest that when teachers make connections between mathematics content to the real world, emphasize how mathematics lessons are applicable to everyday tasks (e.g., how to calculate restaurant tips, store discounts, or grade point average), and/or make math more interesting to students, then African American students' math self-concepts, as well as math task value, may increase. Relevant math instruction also predicted self-concept of math ability and math task value at 8th and 11th grade, converging with previous field experiments linking brief relevance interventions to math task value (Gaspard et al. 2015) as well as brief science relevance interventions to science task value (Hulleman and Harackiewicz 2009).

Previous relevance interventions have suggested that self-generated relevance (e.g., writing a short essay about how math may be relevant for one's future goals) may have differential effects than externally-generated relevance (e.g., being told "you might use mental math to calculate your GPA or figure out tips at restaurants") (Canning and Harackiewicz 2015). While these field experiments are informative, one critical distinction from relevant math instruction is the *duration* of relevance. That is, this study examines students' perceptions of the degree to which their math teacher emphasizes relevance over the course of an academic year, rather than a one-time brief relevance intervention (e.g., Durik and Harackiewicz 2007; Gaspard et al. 2015; Hulleman and Harackiewicz 2009). This may make the distinction between self- and externally-generated relevance less germane, in that (we assume) externally generated relevance from the math teacher is internalized and reflected upon by the student in some way. That is, we assume some reciprocal processes of (external) teacher relevance messaging and students internalizing these relevance messages over time. This distinction makes it difficult to address debates regarding self- versus externally-generated relevance in the literature. However, these findings do suggest that teachers' emphases upon relevance may foster math beliefs among African American students.

From the perspective of expectancy-value theory, these findings support the role of teachers as important factors in the development of students' competence beliefs (Eccles 1994; Wang and Degol 2013). These findings extend previous inquiry by establishing that existing positive impacts of relevant instruction, established with predominantly White samples, also hold for African American students. Given the importance of math beliefs for redressing STEM attainment disparities between African American and White students (e.g., Wang and Degol 2013), these findings are noteworthy from an equity and justice perspective.

However, relevant math instruction had unexpected relationships to concurrent and later math achievement. Eighth grade relevant math instruction was a weak, yet non-significant, predictor of 8th grade math grades and was a significant *negative* predictor of 9th grade Maryland Math standardized achievement scores. Given the associations between relevant math instruction and math beliefs, and between math beliefs and these math achievement variables in our model, coupled with previous literature, these findings were unexpected. Further, these findings cannot be explained away by multicollinear relationships among these predictors; as the math grades were culled from transcript records, we assume that they are accurate measures of school math achievement. The autoregressive relationships between 7th grade math grades (which served as a control variable) and 8th grade math grades ($\beta = .44$) as well as between 8th grade math grades and 9th grade Maryland Math scores ($\beta = .17$) were significant, yet weaker than would be expected. Given that these autoregressive relationships essentially measure the "ranking" of a participant in the same (or, similar) variable over time, this suggests that math achievement over time was unstable among these participants, particularly in the relationships between 8th grade math grades and 9th grade

achievement scores (Kline 2010). In comparing 8th grade math grades to 9th grade achievement scores, it may also be that math grades and standardized achievement scores measure distinct aspects of math achievement, particularly among African American students (who may also experience greater levels of stereotype threat in a standardized testing situation than in day-to-day classroom performance; Wang and Degol 2013). Statistically, the greater levels of instability made it more difficult to identify predictors of standardized Maryland Math achievement (in particular), in comparison to the more stable autoregressive relationships for self-concept of math ability and math task value. These unexpected findings should be examined further in subsequent research.

The second major way that these findings advance the literature is by illuminating how school racial climates, as measured by teacher differential treatment, played a subtle yet pernicious role in African American students' math beliefs and achievement. During 8th and 11th grade, teacher differential treatment did not significantly predict self-concept of math ability or math task value. This was surprising, considering previous research linking perceptions of differential treatment to decreased (domain-general) levels of student motivation, self-efficacy beliefs, and achievement (Benner and Graham 2013; Byrd and Chavous 2011; Cogburn et al. 2011).

However, teacher differential treatment did negatively predict relevant math instruction during 8th grade. Further, teacher differential treatment exhibited a significant negative relationship to students' "downstream" self-concept of math ability and task value, as mediated by relevant instruction. That is, teacher differential treatment corroded the salutary benefits of relevant instruction on students' self-concept of math ability and task value. This converges with previous research that has linked negative school racial climates and discriminatory teacher actions to decreases in academic beliefs, values, and achievement (Eccles and Roeser 2011; Wang and Degol 2013), but also complicates it. Our results show that this relationship may not be direct (i.e., teacher discrimination may not directly affect student behaviors and self-beliefs) as previously understood, but may exert a mediated corrosive function via students' perceptions of teacher instruction. By examining a specific academic domain (i.e., math), these findings also advance the school climate literature, which has focused on linkages between climate and domain-general academic beliefs as well as achievement.

Teachers' differential treatment had a more limited role than hypothesized. Teacher differential treatment may have simply happened less often in the unique setting of MADICS—Prince George's County was more racially and socioeconomically integrated than most other US counties when these MADICS data were collected (Cook et al.

2002; Gutman et al., in press). The item means for the perceived teacher discrimination measure in this study were more than a full point lower (1.4–1.6), and in some cases, half of the reported value of a related measure of perceived peer discrimination among all students (3.12) reported in Benner et al. (2015). This difference suggests that students perceived the racial climate in Prince George's County schools as better than those schools surveyed in AddHealth, a nationally representative sample of students in grades 7–12. This may be due to the fact that students in the Prince George's County school system came from more racially and socioeconomically integrated neighborhoods (Cook et al. 2002).

Another explanation for the more limited impacts of teachers' differential treatment may be student agency. Teacher differential treatment was not invariant over time, which suggests that participants' views of racialized mistreatment from teachers changed as they matured. Inspection of observed item means in Table 1 suggests that students perceived lower levels of differential treatment as they transitioned from middle (wave three) to high school (wave four). Yet, this analytic strategy cannot discern whether this is reflective of "true" lowered levels of teacher differential treatment in high school, whether students' perceptions of teacher differential treatment changed over time, or some combination of both. African American students may learn (by observation or from other students) which teacher(s) discriminate against African American students, and may act to insulate themselves against teachers' discrimination. For example, it is plausible that African American students may disengage from or be less active in classrooms where they witness teacher differential treatment (e.g., teachers calling on African American students less often, disciplining African American students more harshly)—in order to avoid being the target of teachers' differential treatment themselves (McGee and Martin 2011; Milner 2006). This anticipatory behavior on the part of students may minimize reports of teachers' differential treatment, in that students learn to agentically minimize situations where teachers could discriminate against them. In turn, this would attenuate reports of how frequently teachers treat African American students differently. This complex series of events could not be examined with these data, but may frame how African American students resist discrimination in schooling.

Selection effects may also explain this relationship. Teachers who value diversity may have selected into teaching in the more integrated Prince George's County schools. Alternatively, teachers would have positive intergroup contact with diverse students, leading to reductions in teacher prejudice over time—or biased teachers may have voluntarily or involuntarily left the district (Green et al. 1988).

The relevant instruction predictor variable and the self-concept of math ability, task value, and math achievement outcome variables all measured the same academic domain—math—while the teacher differential treatment measure was domain-general. The matching domain-specificity of the relevant instruction measure may have inflated estimates of relationships between relevant instruction, math beliefs, and math achievement, in comparison to the domain-general teacher differential treatment measure. SEM attenuates this type of error bias (Kline 2010), yet, the match or mismatch of academic domains may have inflated or depressed path coefficients in the structural model.

In sum, these findings underscore the importance of relevant math instruction for African American students in a more racially and socioeconomically integrated district. This study suggests practices that may foster African American students' STEM achievement, considering the importance of math in STEM success and African Americans' underrepresentation in STEM fields. Conversely, perceived teacher discrimination appears to corrode students' perceptions of teachers, and, in turn, African American students' math beliefs—suggesting a mediating pathway that was not identified in previous school climate research.

School racial climate is a broad and multifaceted construct. Teacher differential treatment is a key dimension of school racial climate in this literature (e.g., Benner et al. 2015; Byrd and Chavous 2011). Yet, this study could not examine other relevant dimensions of school racial climate, such as peer racial climates (e.g., peer racial/ethnic discrimination). We were also unable to examine school-level perceptions of racial climate or relevant math instruction with MADICS data, due to an insufficient number of students from each of the twenty-three public schools in Prince George's County, Maryland to support multilevel analyses. We were, therefore, unable to examine the nesting of individual perceptions within shared or school-level perceptions of racial climate (or, of relevant instruction); the former an innovation of Benner et al. (2015). The inability to model the clustering of students in MADICS may have resulted in potential bias to standard errors, which may result in inaccurate significance tests. MADICS does not afford the capacity to calculate intraclass correlations or other informative estimates of potential bias to standard errors, due to the complexity of its data structure (i.e., participants attended 23 middle schools and 14 high schools within Prince George's County alone) and the limited number of students sampled per school. On the other hand, the failure to account for clustering may entail quite minimal bias (Heck and Thomas 2015), and a series of studies have used MADICS without adjusting for potential standard error bias (e.g., Byrd and Chavous 2011;

Chavous et al. 2008; Cogburn et al. 2011; Gutman et al., in press; Rapa and Diemer 2016; Wong et al. 2003). These limitations are offset by the affordances provided by MADICS, such as the capacity to longitudinally examine the mediated pattern of relationships between teacher differential treatment and African American students' self-concept of math ability and math task value, which were not examined in Benner et al. (2015). The minimization of measurement error via the use of SEM, inclusion of a lagged achievement control variable, and use of ITCV estimates also collectively enhance confidence in the statistical inferences drawn from these analyses, reducing the likelihood that the failure to account for clustering would result in an inappropriate statistical inference.

Relatedly, a growth modeling approach would complement the insights of this study, by modeling intra- and inter-individual changes in the processes of interest—as well as how changes in students' perceptions of teachers over time may predict growth in math outcomes, in a latent difference score framework (McArdle 2009). However, the 3-year interval between waves three and four, measurement of these teacher and student variables at only these two waves, and some incompatibilities between how teachers were measured and a growth modeling perspective (e.g., perceptions of teacher differential treatment and relevant math instruction are more likely a function of which teacher(s) students are assigned to than a more "fixed" student trait) made growth modeling less suitable for this study. The insights of this study would be complemented by examining growth trajectories with a different data source.

The 3-year interval between waves three and four also precluded testing a cross-lagged panel model (because any cross-lagged effects would likely weaken over such a time span), which would help in disentangling the direction of some hypothesized relationships—in particular, the relationship between relevant math instruction and self-concept of math ability. Self-concept of math ability and teacher practices are dynamic and interactive processes that may reciprocally cause each other, which is consistent with the structural model and reverse causality models tested. The dynamic interplay between teachers' instructional practices and students' beliefs should be more carefully examined with fine-grained microdata (or other data sources that contain multiple and closely spaced measurements of each construct) to fully understand these potentially interwoven processes.

Math achievement was not measured at wave four (11th grade) in MADICS, which limited our ability to link the study's constructs of interest to more distal measures of achievement. However, we were able to link processes of interest to math achievement in 8th and 9th grade by looking at both school grades and standardized test scores,

while controlling for a prior measure of math achievement (7th grade math grade).

This study examined subjective task value without being able to delineate its four specific dimensions (i.e., intrinsic, attainment, utility, or cost value). As one would expect, relationships among expectancy-value constructs may operate differently when examining intrinsic versus utility value, for example (e.g., Durik et al. 2006; Eccles 2005; Trautwein et al. 2012). Future research could build upon the insights of this study by examining whether teacher interactions may differentially predict the four specific facets of subjective task value, such as whether relevant math instruction differentially predicts students' intrinsic value of math or their math utility value.

MADICS did not measure perceived differential treatment *specific to students' math teacher(s)*, instead surveying participants about differential treatment from all teachers. This limited our ability to discern whether the participants were specifically feeling discriminated against by their math teachers, another teacher, or all teachers. Particularly in academic subjects where African Americans have been previously stigmatized, such as math, teachers' discrimination may have a more pronounced impact on achievement due to the status and power they exert over students' academic success (Benner and Graham 2013). Therefore, in order to accurately understand how teacher discriminatory behaviors may influence achievement in math, future research should examine the potentially differential and domain-specific effects of teacher discrimination. That is, discrimination from any teacher likely undermines students' perceptions of their teachers—while discrimination from the math teacher may have more powerful effects on students' math beliefs and achievement. Relatedly, while 36 % of teachers and 67 % of students in Prince George's County identified as African American (Cook et al. 2002), we could not specifically identify or model the race of participants' math teachers in these analyses, which is a limitation that should be addressed in future research.

Conclusion

This study advances our collective understanding of how relevant instruction promotes and differential treatment corrodes African American students' self-concept of math ability, math task value, and achievement over time—while controlling for gender and prior achievement and while estimating the robustness of study inferences. By focusing on a specific academic domain (i.e., math), these findings advance school climate research, which has focused upon linkages between school climate and domain-general academic beliefs and outcomes. In this study, teachers'

differential treatment appeared to corrode students' perceptions of their teachers (i.e., the salutary benefits of relevant instruction), suggesting a mediated relationship among these processes that has been under-examined in previous research. Future research should continue to identify factors and policies that narrow opportunity gaps and foster math achievement and STEM success among underrepresented groups.

References

Adelman, C. (1999). *Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Benner, A. D., Crosnoe, R., & Eccles, J. S. (2015). Schools, peers, and prejudice in adolescence. *Journal of Research on Adolescence, 25*, 173–188.

Benner, A. D., & Graham, S. (2013). The antecedents and consequences of racial/ethnic discrimination during adolescence: Does the source of discrimination matter? *Developmental Psychology, 49*(8), 1602.

Byrd, C. M., & Chavous, T. (2011). Racial identity, school racial climate, and school intrinsic motivation among African American youth: The importance of person–context congruence. *Journal of Research on Adolescence, 21*(4), 849–860.

Canning, E. A., & Harackiewicz, J. M. (2015). Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility–value information. *Motivation Science, 1*(1), 47.

Chavous, T. M. (2005). An intergroup contact-theory framework for evaluating racial climate on predominantly white college campuses. *American Journal of Community Psychology, 36*(3–4), 239–257.

Chavous, T. M., Rivas-Drake, D., Smalls, C., Griffin, T., & Cogburn, C. (2008). Gender matters too: The influence of school racial discrimination and racial identity on academic engagement outcomes among African American adolescents. *Developmental Psychology, 44*, 637–654.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309.

Cogburn, C. D., Chavous, T. M., & Griffin, T. M. (2011). School-based racial and gender discrimination among African American adolescents: Exploring gender variation in frequency and implications for adjustment. *Race and Social Problems, 3*, 25–37.

Cook, T. D., Herman, M. R., Phillips, M., & Settersten, R. A, Jr. (2002). Some ways in which neighborhoods, nuclear families, friendship groups, and schools jointly affect changes in early adolescent development. *Child Development, 73*(4), 1283–1309.

Davis-Kean, P. E., Huesmann, L. R., Jager, J., Collins, W. A., Bates, J. E., & Lansford, J. E. (2008). Changes in the relation of self-efficacy beliefs and behaviors across development. *Child Development, 79*(5), 1257–1269.

Denissen, J. J., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development, 78*(2), 430–447.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.

Diemer, M. A., Mistry, R., Wadsworth, M. E., López, I., & Reimers, F. (2013). Best practices in conceptualizing and measuring social class in psychological research. *ASAP. Analyses of Social Issues and Public Policy, 13*(1), 77–113.

Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology, 99*, 597–610.

Durik, A. M., Vida, M., & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology, 98*(2), 382.

Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly, 18*(4), 585–609.

Eccles, J. S. (2005). Studying gender and ethnic differences in participation in math, physical science, and information technology. *New Directions for Child and Adolescent Development, 2005*(110), 7–14.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Achievement and achievement motivation. In J. T. Spence (Ed.), *Expectancies, values and academic behaviors* (pp. 75–146). San Francisco, CA: WH Freeman.

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence, 21*, 225–241.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*(3), 215–225.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109–132.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Frank, K. A. (2000). Impact of a confounding variable on a regression coefficient. *Sociological Methods & Research, 29*(2), 147–194.

Frank, K. A., Muller, C., Schiller, K. S., Riegle-Crumb, C., Mueller, A. S., Crosnoe, R., & Pearson, J. (2008). The social dynamics of mathematics coursetaking in high school. *American Journal of Sociology, 113*(6), 1645.

Gaspard, H., Dicke, A. L., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B., & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology, 51*(9), 1226.

Graham, S. (1994). Motivation in african americans. *Review of Educational Research*, *64*(1), 55–117.

Green, C. W., Adams, A. M., & Turner, C. W. (1988). Development and validation of the school interracial climate scale. *American Journal of Community Psychology, 16*(2), 241–259.

Gutman, L., Eccles, J. S. & Malanchuk, O. (in press). Making it through adolescence: Developmental trajectories of African American and European American youth. *Society for Research on Child Development Monographs.*

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus.* London: Routledge.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*(4), 880.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410–1412.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*(2), 509–527.

Kline, R. B. (2010). *Principles and practice of structural equation modeling: Methodology in the social sciences* (3rd ed.). New York, NY: Guilford Press.

Marsh, H. W., Trautwein, U., Ludtke, O., Koller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*, 397–416.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577–605.

McGee, E. O., & Martin, D. B. (2011). "You would not believe what I have to go through to prove my intellectual value!": Stereotype management among academically successful Black mathematics and engineering students. *American Educational Research Journal, 48*(6), 1347–1389.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Esychology, 82*(1), 60.

Milner, H. R, I. V. (2006). The promise of Black teachers' success with Black students. *Educational Foundations, 20*, 89–104.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide*. Los Angeles, California: Muthén & Muthén.

Newmann, F. M., & Wehlage, G. G. (1993). Five standards of authentic instruction. *Educational Eeadership, 50*(7), 8–12.

Okeke, N. A., Howard, L. C., Kurtz-Costes, B., & Rowley, S. J. (2009). Academic race stereotypes, academic self-concept, and racial centrality in African American youth. *Journal of Black Psychology, 35*(3), 366–387.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. New York: Cambridge University Press.

Rapa, L. J., & Diemer, M. A. (2016, under review). *Preparation for bias as a moderator in predicting African American youths' motivation and academic achievement.*

Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal, 38*(2), 437–460.

Sáinz, M., & Eccles, J. (2012). Self-concept of computer and math ability: Gender implications across time and within ICT studies. *Journal of Vocational Behavior, 80*(2), 486–499.

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18*(4), 210–222.

Simpkins, S. D., & Davis-Kean, P. E. (2005). The intersection between self-concepts and values: Links between beliefs and choices in high school. *New Directions for Child and Adolescent Development, 2005*(110), 31–47.

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*(1), 70.

Steiner, P., Cook, T. D., Shadish, W., & Clark, M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods, 15*(3), 250–267.

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763.

Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology, 44*(5), 331–349.

Wang, M. T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Developmental Review, 33*(4), 304–340.

Wang, M. T., & Eccles, J. S. (2014). Multilevel predictors of math classroom climate: A comparison study of student and teacher perceptions. *Journal of Research on Adolescence*. doi:10.1111/jora.12153

Watt, H. M. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th-through 11th-grade Australian students. *Child Development, 75*(5), 1556–1574.

Watt, H. M. (2006). The role of motivation in gendered educational and occupational trajectories related to math. *Educational Research and Evaluation, 12*(4), 305–322.

Wigfield, A. (1994). Expectancy–value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*, 49–78.

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*(3), 265–310.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81.

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. *Handbook of Child Psychology and Developmental Science, 3*, 657.

Wong, C. A., Eccles, J. S., & Sameroff, A. (2003). The influence of ethnic discrimination and ethnic identification on African American adolescents' school and socioemotional adjustment. *Journal of Personality, 71*, 1197–1232.

Matthew A. Diemer is an Associate Professor in the Combined Program in Education and Psychology and Educational Studies programs, as well as a Faculty Associate in the Research Center for Group Dynamics within the Institute for Social Research, all at the University of Michigan. His research interests include critical consciousness, social class, and marginalized youths' development.

Aixa D. Marchand is a doctoral student in the Combined Program in Education and Psychology doctoral program at the University of Michigan. Her research interests include racial socialization practices among families of color, as well as how racial identity may serve as a protective factor and help shape youths' academic self-concept.

Sarah E. McKellar is a doctoral student in the Combined Program in Education and Psychology doctoral program at the University of Michigan. Her research interests focus on how social interactions and conditions shape student and teacher motivation, particularly how stereotypes and academic achievement labels influence motivation and how to develop effective interventions in this area.

Oksana Malanchuk is a Research Investigator in the Research Center for Group Dynamics at the University of Michigan. Her research interests have centered on social identities, their development, and their outcomes. Her recent work focuses on occupational aspirations, social identities, and achievement across the lifespan.