Examining How Proactive Management and Culturally Responsive Teaching Relate to Student Behavior: Implications for Measurement and Practice

Kristine E. Larson

Elise T. Pas

Johns Hopkins University, Bloomberg School of Public Health

Catherine P. Bradshaw

University of Virginia, Curry School of Education

Michael S. Rosenberg

State University of New York- New Paltz, School of Education

Norma L. Day-Vines

Johns Hopkins University, School of Education

Larson, K. E., Pas, E. T., Bradshaw, C. P., Rosenberg, M. S., & Day-Vines, N. L. (2018). Examining How Proactive Management and Culturally Responsive Teaching Relate to Student Behavior: Implications for Measurement and Practice. *School Psychology Review, 47*(2), 153–166. https://doi.org/10.17105/SPR-2017-0070.V47-2

Published in *School Psychology Review*

Date Received: April 27, 2017
Date Accepted: January 3, 2018
Associate Editor: Anne Gregory

Authors' Notes

ACKNOWLEDGEMENTS: We thank Katrina Debnam, Jessika Bottiani, Sandra Hardee, and Lana Bates for their assistance in developing the Double Check framework. Additionally, we thank Laurie deBettencourt for her feedback on earlier drafts of this paper. Support for this project comes from the Institute of Education Sciences (R305A150221 and R324A110107) and Spencer Foundation to Catherine Bradshaw.

Correspondence regarding this article should be addressed to Kristine E. Larson, Johns Hopkins University, Bloomberg School of Public Health, 415 N. Washington Street, Office 503, Baltimore, MD 21231, E-mail: klarson3@jhu.edu

Abstract

The discipline gap between White students and students of color has increased demand for teacher training in culturally responsive and behavior management practices. Extant research, however, is inconclusive about how culturally responsive teaching practices relate to student behavior or how to assess using such practices in the classroom. Identifying proactive behavior management and culturally responsive teaching practices that are associated with positive student behavior may inform teacher training and bolster efforts to reduce disparities in behavioral and academic performance. The current study examined the association between student behaviors and the observed use of and teacher self-reported efficacy in using culturally responsive teaching and proactive behavior management practices. Data were collected from 274 teachers in 18 schools. Structural equation modeling indicated a statistically significant association between observations of culturally responsive teaching and proactive behavior management practices, with observed positive student behaviors in classrooms. Implications for measurement and practice are discussed.

Keywords: Positive Behavior Support, School Discipline, Teachers, Prevention, Structural Equation Modeling.

Examining How Proactive Management and Culturally Responsive Teaching Relate to Student Behavior: Implications for Measurement and Practice

Even after decades of research, African American students continue to be disproportionally represented in exclusionary disciplinary actions such as office referrals, suspensions, expulsions, and referrals to the office and school counselors for disruptive behavior (Bryan, Day-Vines, Griffin, Holcomb-McCoy, & Moore-Thomas, 2012; Gregory, Skiba, & Noguera, 2010; Noltemeyer & Mcloughlin, 2010; Vincent, Sprague, & Tobin, 2012; Wallace, Goodkind, Wallace, & Bachman, 2008). Disproportionality refers to a phenomenon whereby students, relative to their proportion in the population, experience overrepresentation or underrepresentation along a specific data point (Bryan et al., 2012). Of particular concern is the over-representation of African American students in discipline data, as research suggests they are three times as likely to get suspended as White students (Losen & Gillespie, 2012) and lose approximately twice as many days of instruction to exclusionary discipline as White students (Vincent et al., 2012). Given that the majority of referrals that lead to exclusionary discipline practices are written by teachers for disruptive behaviors that begin in the classroom, it is argued that disproportionality, and the subsequent school-to-prison pipeline, begins in the classroom and therefore, must be addressed in classrooms.

Although there are well-established benefits of teachers engaging in positive and proactive behavior management strategies to prevent and respond to behavioral infractions (Sugai & Horner, 2002; 2006), research demonstrates that their use does not adequately meet the behavioral needs of all students (Siwatu & Starker, 2010) or close the discipline gap for students of color (Vincent, Swain-Bradway, Tobin, & May, 2011). Specifically, such positive behavior supports reduce the overall use of such exclusionary disciplinary responses (e.g., Bradshaw,

Mitchell, & Leaf, 2010), but students of color are still disproportionally represented. Extant literature exploring teachers' use of culturally responsive practices is currently inconclusive (Bottiani, Larson, Debnam, Bischoff, & Bradshaw, 2017); however, it is hypothesized that using such practices would further and more equitably improve student behavioral outcomes and reduce disproportionality. Furthermore, the best approach to measurement of such practices is unknown. Thus, more research is needed on culturally responsive practices, particularly in relation to student outcomes.

The purpose of the current study was to utilize a multi-method approach to assess teachers' use of both proactive behavior management and culturally responsive teaching practices and to examine the extent to which both were associated with student behaviors observed at the classroom level. We were particularly interested in contrasting teacher self-reported efficacy for engaging in such practices as well as observations conducted within the classroom to determine whether differences existed between teacher beliefs and observed practice. In fact, much of the extant research has relied on self-report measures of culturally responsive teaching (Chu, 2013; Chu & Garcia, 2014; Siwatu, 2007, 2009, 2011; Siwatu & Starker, 2010;); the inclusion of both self-reports and observations of culturally responsive teaching, in addition to social desirability bias, allowed us to advance prior research on this topic and inform our understanding of measurement issues related to culturally responsive teaching. This line of inquiry serves as a critical starting point for the identification of indicators and measurement modalities, as well as promising teacher practices, which can then be used to inform further measurement and intervention research.

Proactive Behavior Management

Student engagement and achievement are associated with teachers' use of evidence-based and proactive classroom management practices (Dunlap, Iovannone, Wilson, Kincaid, & Strain, 2010; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). Examples of such practices include clearly stating expectations, anticipating when students may struggle, and pre-correcting for undesired behaviors (e.g., reminding students before they transition to a new activity what behaviors to engage in as a way to avoid them talking to one another and getting off-task; Dunlap et al., 2010; MacSuga, Simonsen, & Briere, 2012; Pisacreta, Tincani, Connell, & Axelrod, 2011; Simonsen et al., 2008). Research suggests that a broader set of positive instructional and behavioral management practices, including providing students with behavior-contingent praise and increasing opportunities to respond, are also associated with improved student academic behaviors (Dufrene, Lestremau, & Zoder-Martell, 2014; Sutherland, Wehby, & Copeland, 2000; Sutherland, Wehby, & Yoder, 2002). However, most of the research on such management strategies has been conducted with small samples of teachers and among students with specific disabilities. Moreover, school-wide implementation of proactive and positive behavior management practices (e.g., setting clear behavioral expectations and reinforcing the meeting of these expectations) has been shown to reduce disciplinary referrals and suspensions generally (Bradshaw, Mitchell, & Leaf, 2010), but has evidenced a limited impact on disproportionality (Vincent et al., 2011). As such, additional research is needed to examine associations of the use of such practices with student behavior in larger samples of classrooms, including ethnically diverse general education classrooms.

Culturally Responsive Teaching

Culture has been defined as the integrated pattern of human behavior (e.g., thoughts, communication, action, customs, beliefs, values, and instructions) of a racial, ethnic, religious, or

social group (Day-Vines et al., 2007; Leighton, 1982). Some have described culture as a spectrum, ranging from surface (e.g., food, language, dress) to deep (e.g., notions of self, norms, prejudices; Hammond, 2014; Weaver, 1986). Scholars suggest that a "disconnect" between the school and home in terms of surface to deep culture may contribute to disproportionality (Cholewa & West-Olatunji, 2008). Thus, it is hypothesized that educators must understand the association between students' culture and behavior (Cholewa & West-Olatunji, 2008; Hosp & Hosp, 2001; Skiba, Poloni-Staudinger, Gallini, Simmons, & Feggins-Azziz, 2006), and incorporate students' culture into their teaching (Gay, 2002; Ladson-Billings, 1995a; 1995b; Villegas & Lucas, 2002) to improve outcomes of students from ethnic/racial minority backgrounds. This understanding and incorporation of student's culture into the classroom is referred to as *culturally responsive teaching* or *culturally responsive pedagogy*. Scholars generally agree that the basic principles of culturally responsive teaching include having knowledge, dispositions, and skills necessary to teach in a diverse society (Villegas & Lucas, 2002).

Engaging in culturally responsive teaching generally begins with knowledge of culture broadly, teachers' clear understanding of their own and others' cultures, and an ability to connect to their students through this understanding (Gay, 2002; Ladson-Billings, 1995a; 1995b; 2001; Villegas & Lucas, 2002). Strategies to foster this connection may include integrating artifacts that reflect students' interests and using real-world examples and problems to solve during instruction to connect students to their community, national, and global identities (Bouillion & Gomez, 2001). Additionally, in recognizing and responding to variability in how students may prefer to demonstrate knowledge, teachers may provide opportunities for students to co-teach or co-facilitate lessons (Emdin, 2008). Moreover, scholars suggest that teachers may accommodate

different communication styles through varying their own communication (Bondy, Ross, Gallingane & Hambacher, 2008; Day-Vines & Day Hairston, 2005). For example, teachers may use humor to engage students, defuse problems, or set expectations (Bondy et al., 2008) or provide direct commands when asking students to complete a task, as some students respond best to such requests (Day-Vines & Day-Hairston, 2005).

Little empirical research has examined the extent to which culturally responsive teaching strategies are associated with student outcomes (Ahram, Fergus & Noguera, 2011; Bottiani et al., 2017; Moore & Ratchford, 2007; Reglin, Akpo-Sanni & Losike-Sedimo, 2009). The few studies examining the association between culturally responsive teaching strategies and student behavior outcomes have typically used small samples, analyzed data using descriptive statistics, and relied heavily on disciplinary and special education referral data (Ahram et al., 2011; Moore & Ratchford, 2007; Reglin et al., 2009). Although referral data can provide information on shifts in the degree of disproportionality, using such data in isolation does not allow one to parse out whether student behavioral or referral process changes are the mechanism for change. Further, much research lacked a rigorous (e.g., randomized) research design, thus one cannot draw causal conclusions about whether a change is associated with improved teacher practices. Moreover, prior research has relied heavily on teacher self-reports and assessed efficacy rather than actual teacher practices (Chu, 2013; Chu & Garcia, 2014; Siwatu, 2007; 2009; 2011; Siwatu & Starker, 2010). As such, additional research that incorporates observations of both teacher behaviors and student outcomes is needed.

Teacher Self-Efficacy and Social Desirability

Self-efficacy is an individual's belief or confidence in his or her capabilities to execute specific actions or tasks (Bandura, 1977). Interest in teacher self-efficacy began with general

teacher self-efficacy (Dembo & Gibson, 1985) and shifted to focus on more specific areas of teacher self-efficacy, including classroom management (Emmer & Hickman, 1991; Lin, Gorrell, & Taylor, 2002; Woolfolk, Rosoff, & Hoy, 1990) and teaching students from diverse backgrounds (Banks, Dunston, & Foley, 2013; Chu, 2013; Guyton & Wesche, 2005; Pang & Sablan, 1998; Siwatu, 2007; Sorrells, Schaller, & Yang, 2004; Tucker et al., 2005). Collecting self-efficacy data can be a time and cost-efficient method to assess teachers' abilities and a proxy for teacher practices. Several studies have found that self-efficacy in behavior management is associated with reductions in disruptive student behaviors (Reinke et al., 2013) and indicators of disproportionality (e.g., Peters, Kranzler, Algina, Smith, & Daunic, 2014; Reinke, Herman, & Stormont, 2013). Specifically, teachers with a greater sense of self-efficacy in behavior management report externalizing behaviors of students of color to be less problematic and more similar to their White counterparts than those teachers with lower self-efficacy in behavior management (Peters et al., 2014). There is likely a transactional association between efficacy and behavior such that teachers with higher self-efficacy in behavior management engage in improved practices, which results in improved student behavior, and improved behavior leaves teachers feeling more efficacious. Thus, teacher self-efficacy in behavior management, as well as actual use of culturally responsive teaching may play an important role in student outcomes, but further research incorporating efficacy and actual strategy use is needed (Bondy et al., 2007; Brown, 2004).

Measuring self-efficacy has figured prominently in culturally responsive research because *knowledge* and *attitudes* are key characteristics of a culturally responsive teacher (Gay, 2002; Ladson-Billings, 1995a; 1995b; 2001; Villegas & Lucas, 2002). However, research, particularly in the counseling field, suggests that self-reports of multicultural self-efficacy, or the

aggregate attitudes/ beliefs, knowledge, and skills to work with individuals from a variety of cultural (i.e., racial, ethnic, gender, social class, and sexual orientation) groups, are likely to be influenced by teachers' social desire to provide positive reports about themselves (see Constantine, 2001; Katz & Hoyt, 2014; Larson & Bradshaw, 2017; Ohm & Rosen, 2011). An improvement on extant self-report literature would take such social desirability bias into account when measuring self-reports of culturally responsive teaching (Spanierman et al., 2011). Determining whether individuals seek social approval on self-report measures is important to advance the field, as is determining whether those ratings are also related to observations of teacher behaviors.

Current Study

The purpose of this study was to examine teachers' use of proactive behavior management and culturally responsive teaching practices simultaneously to determine whether they were associated with student behaviors observed at the classroom level. We utilized a quantitative, multi-method approach, through which we collected both teacher self-reported efficacy and behavioral observations of proactive behavior management and culturally responsive teaching. We used this approach to address multiple gaps in the culturally responsive teaching literature, which has largely excluded quantitative analysis of how culturally responsive teaching associates with student behavior and has mainly relied on self-reported survey research.

Extant literature suggests that self-efficacy is tied to teacher behaviors (i.e., use of specific teaching strategies; Siwatu, 2009), and the use of proactive behavior management (Simonsen et al., 2008) and culturally responsive teaching practices (Moore & Ratchford, 2007) is associated with higher rates of positive student behavior. Thus, we hypothesized that higher self-reported ratings of self-efficacy and ratings by observers of both culturally responsive

teaching and behavior management would be related to observed student behavior. Additionally, since the literature suggests that there may be an association between social desirability bias and teacher self-reports (see Constantine, 2001; Crowne & Marlowe, 1960; Katz & Hoyt, 2014; Larson & Bradshaw, 2017; Ohm & Rosen, 2011), we controlled for social desirability bias. We also controlled for select teacher characteristics, such as race, role (i.e., general educator vs. special educator), and number of years in the teaching role, because prior research suggests that multicultural self-efficacy is associated with practitioner characteristics (see Constantine, Juby & Liang 2001; Granello & Wheaton, 1998; Larson & Bradshaw, 2017; Neville, Spanierman, & Doan, 2006). Furthermore, we adjusted for years of experience, as teachers who are newer to the field may lack experience and expertise to integrate culturally responsive practices, like other effective behavior management practices, into their teaching (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Clotfelter, Ladd, & Vigdor, 2007; Easton-Brooks & Davis, 2009). Yet, educational research has rarely considered these characteristics when exploring self-reports as well as observations of culturally responsive teaching (Chu & Garcia, 2014; Daunic, Correa, & Reyes-Blanes, 2004; Imler, 2009).

Method

Survey and observational data from a larger study were used to examine the associations among teachers' use of proactive behavior management and culturally responsive teaching practices with observed student behaviors at the classroom level. The sections below provide information about the participants, the procedure for data collection, and the measures used in the current study.

Participants

Self-report and observation data were obtained from 274 teachers in 18 schools (n = 106 elementary school teachers in 9 elementary schools; n = 168 middle school teachers in 9 middle schools); all teachers were participating in an intervention study, however, only the baseline (i.e., pre-intervention) data from that project were used for the current study. All schools were located in one district; they included suburban and urban fringe settings and served a diverse student body (see Table 1).

Table 1

Demographics of Teachers and Schools in Current Study

Teacher Characteristic	N (%) *N* = 274
Age	
20-30	85 (30)
31-40	73 (26)
41- 50	62 (23)
51-60	47 (17)
61+	9 (3)
Gender	
Female	237 (86)
School Level	
Pre-K/ Kindergarten	25 (9)
Elementary	72 (26)
Middle	130 (47)
Multiple Grades	49 (18)
Race/ Ethnicity	
Caucasian/ White	221 (80)
African American/ Black	34 (12)
Hispanic/Latino	4 (1)
Asian/ Pacific Islander	2 (1)
Native American/ American	2 (1)
Other	13 (5)
Teacher Role	
General Educator	237 (86)
Middle School	168 (61)
Years of Teaching in Role	
0-3	71 (26)
4-8	74 (27)
9+	131 (47)
School Demographics	*M (SD)*

Percent of African American Students	34.38 (17.41)
Percent of Hispanic Students	12.68 (6.10)
Percent of Students Receiving FARMs	50.05 (22.66)

Note. FARMs is an abbreviation for free and reduced priced meals. School demographic data are collected annually by the state and made openly accessible and are from the study year.

Across the 18 schools, based on full enrollment data, about 41% of the students were White, one-third (34%) of the students were African American, and about 13% of students were Hispanic. Enrollment data also reflected that approximately half of the students in these schools received free and reduced priced meals (FARMs). Based on observational data, where observers recorded their count of White students, approximately 60% (n = 3129) of the students in the classrooms visited were recorded as White. The sample included only general (86%, n = 237) and special (14%, n = 39) educators responsible for delivering core content (i.e., English, Math, Science, Social Studies) within general education classrooms. Most teachers were women (86%; n = 237) and White (80%; n = 221); approximately 26% (n = 71) of the teachers were in their first three years of teaching.

Procedure

Data were collected as a baseline measure for a larger intervention study of a professional development and coaching intervention targeting culturally responsive teaching practices and positive and proactive behavior management. Teacher participation was voluntary and consent was provided to complete surveys and to allow for external observations to be conducted in the classroom. Self-report surveys were administered via a secure online survey system. Participating teachers received a $10 gift card in appreciation of their time completing the

survey. Among consented teachers, 95% completed the self-report survey and 89% had observations conducted.

Observations of teachers' classrooms were collected by research team-trained observers in 248 classrooms prior to the participants being informed of their intervention status. All observational data collectors received training in four stages: (a) an initial didactic session, (b) on-site practice, (c) on-site inter-observer agreement or reliability, and (d) on-site recalibration. Each data collector was trained until they reached a reliability criterion of 80% (for additional information on the scale and its administration, see Debnam, Pas, Bottiani, Cash, & Bradshaw, 2015; Pas, Cash, O'Brennan, Debnam, & Bradshaw, 2015). Dates for observations were coordinated with administrators and teachers were notified that observers would come to their classrooms at some time during the days selected. The specific times within those prescheduled dates were selected by the research team observers; if there was a test or the teacher present was not the teacher of record, the observer would return to the classroom at a different time to conduct the observation.

Measures

Data were collected via an online self-report survey and from classroom observations. The survey included demographic questions and Likert-type items, with *1 = strongly disagree* to *6 = strongly agree*. Four survey scales were of interest for this study and are described below.

Observational data were collected using the *Assessing School Settings: Interactions of Students and Teachers* (ASSIST; Rusby et al., 2001), which assesses information about various student and teacher behaviors. During a 15-minute observation, ASSIST trained data collectors, who were unaware of the purpose of the research study and teachers' intervention status, tallied the number of teacher and student behaviors observed across several dimensions. At the end of

each 15-minute observation session, data collectors left the teacher's classroom and immediately completed a series of global rating items. Global ratings were scored on a 5-point Likert-type scale, with *0 = Never* and *4 = Almost continuously or Often occurred (6+ times)*. Three of the global rating scales (i.e., *Student Cooperation*, *Teacher Proactive Behavior Management*, and *Teacher Culturally-Responsive Strategies*) were of interest in this study. Together, these data were used to examine the extent to which culturally responsive teaching and proactive behavior management practices were associated with student behaviors. A recent generalizability study (G-study) was conducted of the ASSIST and reported strong reliability of the scores (Abry, Cash, & Bradshaw, 2017).

Student behavior outcome. Student engagement in positive behavior was measured using the *Student Cooperation* scale on the ASSIST; this 7-item scale includes indicators of how often students abided by school rules, norms, and expectations for academic readiness. Sample items include, "Students are focused and engaged" and "Students comply." Higher scale scores reflected greater student cooperation. This scale demonstrated adequate internal validity in prior research (Pas et al., 2015) and internal reliability in the current study ($\alpha = .92$).

Teacher self-efficacy in behavior management. Five items were used to measure teachers' self-reported ability to manage challenging student behaviors. Three items on the measure originated from the *Teacher Efficacy Scale* (Hoy & Woolfolk, 1993) and two additional questions were added from previous research (Pas, Bradshaw, & Hershfeldt, 2012). Example items are "If a student becomes disruptive and noisy, I feel assured that I now have some techniques to redirect him/her quickly" and "I can effectively work with deviant or disruptive students." This five-item measure demonstrated adequate internal consistency in previous

research (Pas et al., 2012) and in the current study (α = .80). Higher scale scores indicated greater teacher self-efficacy to manage behavior problems.

Observed proactive behavior management. The *Proactive Behavior Management* subscale of the ASSIST was comprised of 4 items and assessed how often teachers implemented practices that prevented or represented a positive response to behavioral disruptions (α = .66). Sample items include, "Teacher gives clear instructions and directives to students," "Teacher is consistent, even-handed, and firm when necessary," "The teacher clearly explains learning objectives prior to and/or during the lesson through summary or re-orientation statements," and "Teacher praised students for specific behaviors or using social skills." Higher scale scores reflected greater use of proactive behavior management strategies.

Self-efficacy in culturally responsive teaching. The *Multicultural Efficacy Scale* (MES; Guyton & Wesche, 2005) is a 14-item scale that assessed one's efficacy to implement skills needed to teach in multicultural settings (α = .81). Sample items include, "I can adapt instructional methods to meet the needs of learners from diverse groups," "I can help students work through problem situations caused by stereotypical and/or prejudicial attitudes," and "I can develop activities that increase the self-confidence of diverse students." Higher scale scores indicated greater self-efficacy in culturally responsive teaching.

Observed use of culturally responsive teaching. The *Teacher Culturally-Responsive Strategies* subscale of the ASSIST assessed the use of culturally responsive teaching practices in the classroom (7-item α = .73). The seven items on this scale were: "Teacher connects lessons to real world examples," "Teacher integrates cultural artifacts reflective of students' interests into learning activities," "Teacher engages in storytelling or sharing," "Teacher uses positive humor to engage students or defuse problems," "Teacher gives students opportunities to co-teach or co-

facilitate learning," "Teacher gives direct commands," and "Teacher employs rhythm or 'call and response' instructional strategies." Higher scale scores reflect greater observed use of culturally responsive teaching strategies.

Social desirability. The *Marlowe-Crowne Scale of Social Desirability* (Crowne & Marlowe, 1960) is one of the most widely-used measures of social desirability; it measures the extent to which survey respondents answer questions in a manner that they believe will be viewed favorably by others and that individuals are likely not *always* able to agree or disagree (α = .62). Items include, "I am always courteous, even to people who are disagreeable" and "I never hesitate to go out of my way to help someone in trouble." The original Marlowe-Crowne (1960) measure included 33 items and had an internal consistency of .88 and a test-retest correlation of .89 (Crowne-Marlowe, 1960). Shortened forms of the Marlowe-Crowne scale have been shown to be reliable (Ballard, 1992; Fischer & Fick, 1993; Strahan & Gerbasi, 1972) with a Kuder-Richardson coefficient of .62 (Strahan & Gerbasi, 1972). Higher scale scores indicated a higher degree of social desirability bias or a higher degree to which the respondent seeks social approval from others. In the context of other self-report scales, this measure would indicate the extent to which a person may have answered in a socially desirable manner, and thus serves as a potentially important control variable in the current study (Larson & Bradshaw, 2017).

Teacher demographics. Teachers provided self-reported basic demographic data on the online survey. Specifically, teachers reported their ethnic/ racial group (i.e., Native American/ American Indian, Asian/Pacific Islander, Black/African American, White, Hispanic/Latino, Native Hawaiian or Other Pacific Islander, and Other). In addition, teachers reported their role in the school as either a general or special educator. Lastly, teachers reported the number of years they had been teaching in their role. Since teachers are believed to become more effective with

experience (Boyd et al., 2006; Clotfelter et al., 2007; Easton-Brooks & Davis, 2009), we included a variable indicating the extent to which a teacher was new in his or her career (0 = 0-3 years, 1 = 4 or more years).

Analyses

We used structural equation modelling (SEM) using M*plus* version 7.3 (Muthén & Muthén, 2013) to examine our research aims (see Figure 1).

Figure 1. Structural Equation Model Displaying Standardized and Statistically Significant Associations

SEM allowed simultaneous testing of the association and directionality among the latent variables of interest, which are presumed to reflect a construct that is not directly observable (Kline, 2011; Schreiber, Nora, Stage, Barlow & King, 2006; Teo, 2010). Using latent variables in SEM allows for an entire model based on an a priori theory to be estimated and generates an accurate estimate of measurement error (Kline, 2011; Schreiber et al. 2006). Specifically, the error variance in SEM is estimated for the entire model, thereby freeing the latent variables themselves from measurement error (Kline, 2011; Schreiber et al. 2006). SEM also provides information about the association among the variables and the indicators that comprised each latent variable and scale (Kline, 2011). The items reported for each of the measures described above were used to create a latent construct for each scale. Finally, when modeling scales as latent, using the items as indicators, and utilizing the maximum likelihood method, teachers with data on at least some of the items for each construct can be retained. Despite this, we performed a Missing Data Value Analysis, the findings of which are presented in the Results section.

Prior to analyzing the structural equation model, a confirmatory factor analysis (CFA) was examined to replicate earlier measurement work conducted on the ASSIST (e.g., Debnam et al., 2015; Pas et al., 2015). In addition, demographic variables, including teacher role (1 =

general educator, 0 = special educator), experience in role (0 = 0-3 years, 1 = 4 or more years, as an indicator of later career), gender (0 = male, 1 = female), and race (0 = Non-White, 1 = White) were included as manifest (i.e., observed, non-latent) covariates in the model.

Fit indices were examined to determine the fit between the SEM models and the observed data. Specifically, the root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis Index (TFI), and weighted root mean residual (WRMR) were computed to determine model fit (Kline, 2011) for both the CFA and the SEM. Fit indices for the CFA indicated that the items adequately loaded onto their respective a priori scales (RMSEA = .022, CFI = .968, TLI = .966, WRMR = 1.196). Prior research suggests the following values for adequate model fit: RMSEA <.06, CFI > .95, and TFI > .95, and WRMR is close to 1.0 (Hu & Bentler, 1999; Kline, 2011; Schreiber et al., 2006). We took into consideration that survey and observation items were ordinal by using weighted least squares means and variance adjusted (WLSMV) estimation. Finally, given that teachers were clustered into 18 schools, we used the type = complex statement and group-mean centered all variables to ensure that parameter, standard error, and *p*–value estimates accounted for this clustering (Stapleton, McNeish, & Yang, 2016). The range and means of scale scores on the observations were normally distributed and utilized the full range of possible responses on the 5-point Likert scale (see Table 2).

Table 2

Descriptive Statistics of Measures and Correlations

Measure	Range	*M* (SD)
Teacher Self-Reported Measures		
Social Desirability Bias (SD)	2.90 - 5.70	4.46 (0.56)
Multicultural Efficacy Scale (MES)	3.29 - 5.86	4.62 (0.47)
Teacher Efficacy in Behavior Management (TEBM)	2.20 - 6.00	4.51 (0.69)

Observational Measures		
Teacher Proactive Behavior Management (TPBM)	1.25 - 4.00	3.06 (0.61)
Positive Student Behavior (SBOP)	1.00 - 4.00	3.24 (0.66)
Culturally Responsive Teaching Strategies – Observed (CRTO)	0.00 - 4.00	1.21 (0.92)

Measure	MES	TEBM	TPBM	SBOP
Teacher Efficacy in Behavior Management (TEBM)	.51**			
Teacher Proactive Behavior Management (TPBM)	-.07	-.05		
Positive Student Behavior (SBOP)	-.07	-.02	.74**	
Culturally Responsive Teaching–Observed (CRTO)	-.00	.01	.38**	.47**

Note. The Teacher Self-Reported Measures were on a 6-point Likert Scale, with possible scale scores ranging from 1-6. The observational measures were on a 5-point Likert Scale, with possible scale scores ranging from 0-4. ** $p<.01$.

Results

The section below summarizes the findings from the analysis regarding the association among teachers' use of proactive behavior management and culturally responsive teaching practices with observed student behaviors at the classroom level. A description of the analysis of missing data is provided followed by results from the SEM analysis.

Missing Data

An analysis of missing data was conducted to evaluate the amount, distribution, and pattern of missing data. SPSS Missing Values Analysis (MVA) was used to highlight patterns of missing values (Tabachnick & Fidell, 2013). Data were present for more than 95% of the dataset. Specifically, missing data on scales ranged from 0-10% ($M = 4.90$; $SD = 9.24$). The highest percentage of missing data was from observations in the classrooms. Specifically, about 10% ($n = 26$) of data were missing for classroom observation variables (e.g., culturally responsive

teaching, student cooperation, and teacher proactive behavior management). One teacher (0.4%) was missing data on the teacher efficacy for behavior management scale. Teachers were not missing data on the multicultural efficacy scale or social desirability scales. Additionally, data collectors observed transitions in only 35% of classrooms in the sample and one of the items on the student cooperation scale asked about transitions; as such, 65% of data was considered missing for on item on the student cooperation scale. Little's test of Missing Completely at Random, however, was not statistically significant for any value ($\chi = 317.98$, $p > .05$), thereby suggesting that missing data were missing completely at random (Little, 1988). As such, most data were present ($\geq 90\%$) for the majority of the variables and handling such a small amount of missing data using any missing data techniques are not necessary and would yield similar results (Tabachnick & Fidell, 2013).

Structural Equation Model

Positive student behavior. We fit a model to test our hypothesis that higher ratings of student behavior would be provided by observers in classrooms where teachers reported higher self-efficacy in both behavior management and culturally responsive teaching and were observed using culturally responsive teaching and proactive behavior management practices. Fit indices indicated that the data adequately fit this model (RMSEA = .022, CFI = .968, TLI = .966, WRMR = 1.196). The SEM results indicated that observations of culturally responsive teaching (see Table 3) were statistically significant and positively associated with the student behavior outcome, such that a one-standard deviation change in culturally responsive teaching strategies was associated with a 0.12 point increase in the observer ratings of student behavior.

Table 3

Standardized Parameter Estimates for Structural Model (n = 274)

	Coefficient	*SE*	*p*-value		
Predicting Positive Student Behavior (SBOP)					
Behavior Management Measures					
Teacher Self-Efficacy in Behavior Management (TEBM)	0.08	0.05	.10		
Teacher Proactive Behavior Management-Observed (TPBM)	0.84***	0.04	.00		
Culturally Responsive Teaching Measures					
Multicultural Self- Efficacy Scale (MES)	-0.03	0.04	.43		
Culturally Responsive Teaching- Observed (CRTO)	0.12*	0.06	.04		
Standardized Covariates	TEBM	TPBM	MES	CRTO	
Social Desirability Bias	.57***	.00	.46***	.07	
General Educator	-.15**	-.04	-.06	.15*	
4+ Years Teaching in Role	-.04	.01	.03	.08	
Teacher is White	-.06	-.07	-.06	.01	
Teacher is Female	.01	.05	-.13*	.09	
Standardized Covariances	TEBM	TPBM	MES	CRTO	SBOP
TEBM		-.12*	.57***	.01	
TPBM			-.18*	.60***	
MES				.01	
Interclass Correlations (ICC)	0.10	0.00	0.10	0.01	0.20

Note. Model displays standardized regression coefficients loadings. *$p < .05$; ** $p < .01$; *** $p < .001$.

Observations of proactive behavior management were also statistically significant and positively associated with the student behavior outcome. Specifically, a one-standard deviation increase in proactive behavior management strategies was associated with 0.84 point increase in the student behavior outcome. Although both observational measures of teacher strategies were associated with positive student behavior, teacher self-efficacy of behavior management and teacher self-efficacy of culturally responsive teaching were not (See Table 3).

Association between Measures

Self-reports and observations of constructs. The model also allowed for examination of covariance between the self-efficacy and observational measures. Findings indicated that there were statistically significant associations between the two observations of teacher strategy use ($\psi = .60$, $p < .001$) and between the two self-efficacy scales ($\psi = .57$, $p < .001$). Across measures (i.e., between efficacy and observed use of strategies) observations of teacher proactive behavior management were negatively associated with teacher self-efficacy of (behavior management $\psi = -0.12$) and self-efficacy of culturally responsive teaching ($\psi = -0.18$; $ps < .05$; see Table 3).

Social desirability bias. Social desirability was modeled as a latent construct for the purpose of controlling for the possibility that measures could be influenced by such bias. Results indicated that social desirability was statistically significant and positively associated with the self-efficacy reports of culturally responsive teaching ($\psi = .46$, $p < .001$) and teacher efficacy in behavior management ($\psi = .57$, $p < .001$). In other words, higher ratings of teacher self-efficacy for culturally responsive teaching was related to higher social desirability bias, highlighting the importance of including this variable in the study. In contrast, observation measures were not associated with self-reported social desirability (See Table 3).

Teacher characteristics. Several teacher characteristics were statistically significantly associated with self-efficacy and observation measures. Specifically, general educators reported less self-efficacy in behavior management than special educators ($\psi = -.15$, $p < .01$) and were observed to use greater culturally-responsive teaching strategies ($\psi = .15$, $p < .01$). Moreover, females reported lower self-efficacy related to culturally responsive teaching than males ($\psi = -.13$, $p < .01$).

Discussion

The purpose of this study was to determine whether teachers' self-efficacy and observations of culturally responsive teaching and proactive behavior management were associated with observed student behavior at the classroom level. We hypothesized that teachers' use of culturally responsive teaching and proactive behavior management practice, as rated by external observers, together with higher teacher-reported self-efficacy in using such practices would be associated with more favorable student behaviors, even after controlling for teacher social desirability and teacher characteristics. Taken together, the SEM results indicated partial support for our hypothesized model; our analyses indicated that observations of teachers' use of both culturally responsive and proactive behavior management practices were associated with observed positive student behavior, whereas teachers' self-efficacy regarding their use of such skills was not. Prior research suggests that proactive behavior management is important to positive student behavior (Dunlap et al., 2010; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). In fact, when observers in this study noted higher frequencies of proactive behavior management practices (e.g., providing clear instructions, clearly explaining learning objectives, and praising students for specific behaviors), they also observed greater student cooperation and engagement in expected behaviors. Similarly, the teachers' use of culturally responsive teaching practices (e.g., lessons to the real-world, using positive humor, and employing rhythm call and response) was also related to more positive ratings of observed student behavior. Using Cohen's (1992) magnitude interpretation, our statistically significant results suggest that culturally responsive teaching has a small association with positive student behavior, over and above effect of proactive behavior management. It is worth noting, however, that the association between observed proactive behavior management and positive student behavior was considerably larger than the association between culturally responsive teaching and positive student behavior.

Further, the average scores on the culturally responsive teaching practices scale were somewhat low (i.e., 1.21 on a scale up to 4), which indicates additional room for growth on this skill. Nonetheless, the relative magnitude of the culturally responsive teaching finding is a novel contribution of this study.

On the other hand, teachers' self-efficacy of culturally responsive teaching and behavior management was not positively associated with the observed student outcomes. These results were unexpected considering the number of studies that use self-efficacy measures to assess aspects of culturally responsive teaching (Chu, 2013; Chu & Garcia, 2014; Siwatu, 2007, 2009; Siwatu & Starker, 2010), as well as other teacher-efficacy related research dating back to the mid 1980's (Dembo & Gibson, 1985; Emmer & Hickman, 1991; Lin, Gorrell, & Taylor, 2002; Midgley, Feldlaufer, & Eccles, 1989; Riggs & Enochs, 1990; Soodak, & Podell, 1993; Soodak, Podell, & Lehman, 1998; Starko & Schack, 1989; Woolfolk & Hoy, 1990). Teachers' self-efficacy to engage in behavior management and culturally responsive teaching was negatively associated with observed teacher use of proactive behavior management. An earlier study found inconsistent associations (i.e., some desired and some undesired) with observations of teacher strategies and self-reports of efficacy (see Debnam et al., 2015). The findings of the current study suggest that self-reported self-efficacy measures may not be the most valid indicators of actual use of culturally responsive practices and that, optimally, future research would integrate direct observations of teacher practices and student behavior in addition to self-report measures of actual use, rather than self-reported self-efficacy to implement practices.

We also considered a variety of teacher-level demographic variables as possible covariates. For example, we found that teacher race was not a statistically significant predictor of efficacy or the observation measures of behavior management or culturally responsive teaching.

Although the lack of an association between teacher race and student behavior is consistent with prior research (Pas & Bradshaw, 2014), the lack of an association for race and the culturally-relevant outcomes was inconsistent with prior research. In fact, there is considerable interest in policies and programs that aim to increase the number of racial and ethnic minority teachers as an approach for reducing disparities, based on such emerging research suggesting that exposure to same-race teachers is associated with reduced rates of exclusionary discipline for Black students (Lindsay & Hart, 2017). Similarly, we were surprised that years of experience was not a statistically significant predictor of observed or self-reported efficacy in behavior management or culturally responsive teaching, as prior work has suggested that early career teachers may not be as adept at using these approaches are their more experienced peers (Boyd et al., 2006; Clotfelter, Ladd, & Vigdor, 2007; Easton-Brooks & Davis, 2009). This finding suggests that other variables in the model may have accounted for these associations; however, additional research is needed to better understand whether there is a potential link between experience and behavior management or culturally responsive teaching.

Strengths and Limitations

There are several strengths of the current study, including the use of a multi-method approach that controls for potential confounding variables, such as social desirability and teacher characteristics. The relatively high racial and socioeconomic diversity of the student sample was also a strength, as was the use of SEM to simultaneously model associations between all constructs of interest. Nonetheless, there are some limitations to consider. For example, the data were cross-sectional and thus we cannot make causal assumptions. Moreover, regarding measurement, we used slightly abbreviated self-report scales from the original measures to

reduce participant burden and to cover a wide range of constructs, and this resulted in reduced internal consistency. This issue may be particularly pronounced for the social desirability scale.

Furthermore, we only conducted one 15-minute observation on the ASSIST for each teacher. Although multiple observations are recommended when it comes to observational assessment of teacher practices (Cantrel & Kane, 2013), it was not feasible to do so in the current study. This is a limitation of the current research, however, the collection of multiple waves of observational data is also not likely feasible for school personnel in practice and thus this research may guide feedback to use in practice with teachers to improve their classroom practices. Moreover, it is challenging to capture all of the strategies to address "culture" in a single observation measure, which further illustrates the need for additional measure development and testing. Taken together, these findings highlight the importance of multi-method approaches to assessing teacher and student behaviors. Future research should incorporate multiple observations which may be able to capture a broader range of teacher and student behaviors as well as ensure the greatest reliability. Another potential concern is that both the observational indicators were assessed by the same observers using the ASSIST, and therefore have shared method variance which may impact construct validity. Thus, the inclusion of other data sources, such as school records of discipline and student reports would further enhance our understanding of the pattern of findings across multiple data sources as well as provide additional insight into disproportionality. It would also be helpful to consider the outcomes for disproportionality more specifically, and include specific measures of both race and socioeconomic status in relation to disproportionality; this information was not available in the current study.

The district, principals, and teachers all volunteered to participate in the project, which may reflect some level of selection bias toward schools and participants that are interested in contributing to the research on issues related to culturally responsive teaching. Finally, the participating teachers only worked in elementary and middle schools. Therefore, additional work is needed on high school teachers and with larger samples allowing for comparisons across the different school levels.

Conclusions and Implications

Taken together, these findings suggest statistically significant associations between observation of culturally responsive teaching and proactive behavior management and positive student behaviors; however, the findings seem to be most salient for observations rather than self-reported indicators of efficacy in behavior management and culturally responsive practices. Although this study is non-experimental, it appears possible that the cumulative use of culturally response teaching strategies could prove promising in helping to address the exclusionary discipline crisis. These practices include making the curriculum relevant to students (e.g., connecting to real-world examples and incorporating cultural artifacts), varying the way in which students' engage and display understanding (e.g. co-teaching) or the way in which the teacher communicates (e.g., using humor, providing direct commands). Determining whether the use of such strategies is associated with a reduction in disparities in exclusionary discipline actions between African American students and White students remains an important next step in this line of research, and certainly an area needing a stronger empirical base (Bottiani et al., 2017).

The current findings also highlight the importance of incorporating both self-reports of efficacy and observed teacher practice, as well as other data sources (Chu, 2013; Chu & Garcia, 2014; Imler, 2009; Siwatu, 2007, 2009; Siwatu & Starker, 2010), when conducting studies

related to culturally responsive teaching. Similarly, the observed link between both of the self-report measures (i.e., teacher efficacy in behavior management and multicultural efficacy) and the social desirability measure, but neither of the observational measures (i.e., from the ASSIST), highlights the importance of accounting for the social desirability in future self-report research (Larson & Bradshaw, 2017).

Regarding implications for practice, school psychologists may want to collect data using both self-report and observational data in the classroom regarding proactive behavior management and culturally responsive teaching, as measured here. Such data can inform professional development efforts to improve a range of student outcomes. School psychologists may also want to provide coaching or consultation support to individual teachers to further augment school-wide professional development activities in these areas. The use of observational approaches, rather than relying on self-reports, may be best for identifying teachers most in need of support. For example, school psychologists could, either independently or in conjunction with administrators and behavior support teams, administer items from the ASSIST culturally responsive practices measure when conducting classroom observations as well as “walk throughs.” Such observational data could also be helpful in evaluating the broader impact of programs and professional development efforts aimed at reducing disproportionality and the discipline gap.

References

Abry, T., Cash, A. H., & Bradshaw, C. P. (2017). Using generalizability theory to examine sources of variance in observed behaviors within high school classrooms. Manuscript submitted for publication.

Ahram, R., Fergus, E., & Noguera, P. (2011). Addressing racial/ethnic disproportionality in special education: Case studies of suburban school districts. *Teachers College Record, 113*, 2233-2266. doi:10.1037/e716862007-001

Ballard, R. (1992). Short forms of the Marlowe-Crowne social desirability scale. *Psychological Reports, 71*, 1155-1160. doi:10.2466/pr0.1992.71.3f.1155

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215. doi:10.1037/0033-295x.84.2.191

Banks, J., Dunston, Y. L., & Foley, T. E. (2013). Teacher efficacy as a conduit for enhancing attitudes toward teaching reading to African-American students. *Multicultural Perspectives, 15*, 19-26. doi:10.1080/15210960.2013.754286

Bondy, E., Ross, D. D., Gallingane, C., & Hambacher, E. (2007). Creating environments of success and resilience: Culturally responsive classroom management and more. *Urban Education, 42*, 326-348. doi:10.1177/0042085907303406

Bottiani, J. H., Larson, K. E., Debnam, K., Bischoff, C., & Bradshaw, C. P. (2017). Promoting educators' use of culturally responsive practices: A systematic review of in-service interventions. *Journal of Teacher Education*. doi:10.1177/0022487117722553

Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2016). Inequality in Black and White high school students' perceptions of school support: An examination of race in context. *Journal of Youth and Adolescence, 45*(6), 1176-1191. doi:10.1007/s10964-015-0411-0

Bouillion, L. M., & Gomez, L. M. (2001). Connecting school and community with science learning: Real world problems and school–community partnerships as contextual scaffolds. *Journal of Research in Science Teaching, 38*(8), 878-898. doi:10.1002/tea.1037.abs

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation & Policy Analysis, 31*(4), 416-440. doi:10.3102/0162373709353129

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions, 12*, 133-148. doi:10.1177/1098300709334798

Brown, D. F. (2004). Urban teachers' professed classroom management strategies reflections of culturally responsive teaching. *Urban Education, 39*, 266-289. doi: 10.1177/0042085904263258

Bryan, J., Day-Vines, N. L., Griffin, D., & Moore-Thomas, C. (2012). The disproportionality dilemma: Patterns of teacher referrals to school counselors for disruptive behavior. *Journal of Counseling & Development, 90*(2), 177-190. doi:10.1111/j.1556-6676.2012.00023.x

Cantrel, S., & Kane, T. (2013). Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three year study. Research Paper. MET Project. *Bill & Melinda Gates Foundation*. doi:10.3886/icpsr34771.v1

Cholewa, B., & West-Olatunji, C. (2008). Exploring the relationship among cultural discontinuity, psychological distress, and academic outcomes with low-income, culturally

diverse students. *Professional School Counseling, 12*(1), 54-61. doi:10.5330/psc.n.2010-12.54

Chu, S. Y. (2013). Teacher efficacy beliefs toward serving culturally and linguistically diverse students in special education: Implications of a pilot study. *Education and Urban Society, 45*, 385-410. doi:10.1177/0013124511409401

Chu, S. Y. & Garcia, S. (2014). Culturally responsive teaching efficacy beliefs of in-service special education teachers. *Remedial and Special Education, 35*, 218-232. doi: 10.1177/0741932513520511

Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682. doi:10.1016/j.econedurev.2007.10.002

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159. doi:10.1037/0033-2909.112.1.155

Constantine, M. G. (2001). Multiculturally-focused counseling supervision: Its relationship to trainees' multicultural counseling self-efficacy. *The Clinical Supervisor, 20*, 87-98. doi:10.1300/J001v20n01_07

Constantine, M. G., Juby, H. L., & Liang, J. C. (2001). Examining multicultural counseling competence and race-related attitudes among White marital and family therapists. *Journal of Marital and Family Therapy, 27*, 353-362. doi:10.1111/j.1752-0606.2001.tb00330.x

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354. doi:10.1037/h0047358

Daunic, A. P., Correa, V. I., & Reyes-Blanes, M. E. (2004). Teacher preparation for culturally diverse classrooms: Performance-based assessment of beginning teachers. *Teacher Education and Special Education, 27,* 105-118. doi:10.1177/088840640402700203

Day-Vines, N. L., & Day-Hairston, B. O. (2005). Culturally congruent strategies for addressing the behavioral needs of urban, African American male adolescents. *Professional School Counseling*, 236-243.

Day-Vines, N. L., Wood, S. M., Grothaus, T., Craigen, L., Holman, A., Dotson-Blake, K., & Douglass, M. J. (2007). Broaching the subjects of race, ethnicity, and culture during the counseling process. *Journal of Counseling & Development, 85*(4), 401-409. doi:10.1002/j.1556-6678.2007.tb00608.x

Debnam, K. J., Pas, E. T., Bottiani, J., Cash, A., & Bradshaw, C. P. (2015). An examination of the association between observed and self-reported culturally responsive teaching practices. *Psychology in the Schools, 52,* 533-548. doi:10.1002/pits.21845

Dembo, M. H., & Gibson, S. (1985). Teachers' sense of efficacy: An important factor in school improvement. *The Elementary School Journal, 86*(2), 173-184. doi: 10.1086/461441

Dufrene, B. A., Lestremau, L., & Zoder-Martell, K. (2014). Direct behavioral consultation: Effects on teachers' praise and student disruptive behavior. *Psychology in the Schools, 51*, 567-580. doi:10.1002/pits.21768

Dunlap, G., Iovannone, R., Wilson, K. J., Kincaid, D. K., & Strain, P. (2010). Prevent-teach reinforce: A standardized model of school-based behavioral intervention. *Journal of Positive Behavior Interventions, 12*, 9–22. doi: 10.1177/1098300708330880

Easton-Brooks, D., & Davis, A. (2009). Teacher qualification and the achievement gap in early primary grades. *Education Policy Analysis Archives, 17*(15). Retrieved from http://epaa.asu.edu/ojs/article/view/17 doi:10.14507/epaa.v17n15.2009

Emdin, C. (2008). The three C's for urban science education. *Phi Delta Kappan, 89*(10), 772-775. doi: 10.1177/003172170808901018

Emmer, E. T., & Hickman, J. (1991). Teacher efficacy in classroom management and discipline. *Educational and psychological measurement, 51*, 755-765. doi:10.1177/0013164491513027

Fischer, D. G., & Fick, C. (1993). Measuring social desirability: Short forms of the Marlowe-Crowne social desirability scale. *Educational and Psychological Measurement, 53*, 417-424. doi:10.1177/0013164493053002011

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education, 53*, 106-116. doi:10.1177/0022487102053002003

Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice*. New York, NY: Teachers College Press.

Granello, D., & Wheaton, J. E. (1998). Self-perceived multicultural competencies of African American and European American vocational rehabilitation counselors. *Rehabilitation Counseling Bulletin, 42*, 2-15.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher, 39*, 59-68. doi:10.3102/0013189x09357621

Guyton, E. M., & Wesche, M. V. (2005). The Multicultural Efficacy Scale: Development, item selection, and reliability. *Multicultural Perspectives, 7,* 21-29. doi:10.1207/s15327892mcp0704_4

Hammond, Z. (2014). *Culturally responsive teaching and the brain: Promoting authentic engagement and rigor among culturally and linguistically diverse students*. Corwin Press.

Hosp, J. L., & Hosp, M. K. (2001). Behavior differences between African-American and Caucasian students: Issues for assessment and intervention. *Education and Treatment of Children, 24,* 336–350.

Hoy, W. K., & Woolfolk, A. E. (1993). Teachers' sense of efficacy and the organizational health of schools. *The Elementary School Journal, 93,* 355-372. doi: 10.1086/461729

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1-55. doi:10.1080/10705519909540118

Imler, S. J. (2009). Becoming culturally responsive: A need for preservice teacher candidates. *Teacher Education and Practice, 22* (3), 351-367.

Katz, A. D., & Hoyt, W. T. (2014). The influence of multicultural counseling competence and anti-Black prejudice on therapists' outcome expectancies. *Journal of Counseling Psychology, 61*(2), 1-7. doi:10.1037/a0036134

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Ladson-Billings, G. (1995a). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32,* 465-491. doi:10.3102/00028312032003465

Ladson-Billings, G. (1995b). But that's just good teaching! The case for culturally relevant pedagogy. *Theory into Practice, 34,*159-165. doi:10.4135/9781452219639.n8

Ladson-Billings, G. (2001). *Crossing over to Canaan: The journey of new teachers in diverse classrooms*. San Francisco, CA: Jossey-Bass.

Larson, K. E., & Bradshaw, C. P. (2017). Cultural competence and social desirability among practitioners: A systematic review of the literature. *Children and Youth Services Review, 76, 100–111*. doi:10.1016/j.childyouth.2017.02.034

Leighton, D. C. (1982). As I knew them: Navajo women in 1940. *American Indian Quarterly*, 34-51. doi:10.2307/1183595

Lin, H., Gorrell, J., & Taylor, J. (2002). Influence of culture and education on U.S. and Taiwan preservice teachers' efficacy beliefs. *Journal of Educational Research, 96*(1), 37-46. doi:10.1080/00220670209598789

Lindsay, C. A., & Hart, C. M. (2017). Exposure to same-race teachers and student disciplinary outcomes for black students in North Carolina. *Educational Evaluation and Policy Analysis*, 39(3), 485–510. doi:10.3102/0162373717693109

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202. doi:10.2307/2290157

Losen, D. J., & Gillespie, J. (2012). Opportunities suspended: The disparate impact of disciplinary exclusion from school. Retrieved from: http://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/upcoming-ccrr-research/

MacSuga-Gage, A. S., Simonsen, B., & Briere, D. E. (2012). Effective Teaching Practices: Effective Teaching Practices that Promote a Positive Classroom Environment. *Beyond Behavior, 22*(1), 14–22. doi:10.1177/107429561202200104

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology, 81*(2), 247-258. doi:10.1037/0022-0663.81.2.247

Moore, E., & Ratchford, V. F. (2007). Decreasing discipline referrals for African American males in middle school. *AASA Journal of Scholarship & Practice, 4*(2), 20-24.

Muthén, L. K., & Muthén, B. O. (2013). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.

Neville, H., Spanierman, L., & Doan, B. (2006). Exploring the association between color-blind racial ideology and multicultural counseling competencies. *Cultural Diversity and Ethnic Minority Psychology, 12*, 275-290. doi:10.1037/1099-9809.12.2.275

Noltemeyer, A., & Mcloughlin, C. S. (2010). Patterns of exclusionary discipline by school typology, ethnicity, and their interaction. *Penn GSE Perspectives on Urban Education, 7*, 27-40.

Ohm, R., & Rosen, L. (2011). Psychometric evaluation of the Perceived Prejudice in Health Care Scale–Modified (PPHC-M) among baccalaureate student nurses. *Journal of Transcultural Nursing, 22*, 282-289. doi:10.1177/1043659611404430

Pang, V., & Sablan, V. A. (1998). Teacher efficacy: How do teachers feel about their abilities to teach African American students? In M. E. Dilworth (Ed.), *Being responsive to cultural differences: How teachers learn* (pp. 39-58). Thousand Oaks, CA: Corwin Press.

Pas, E., & Bradshaw, C. P., (2014). What affects teacher ratings of student behaviors? The potential influence of teachers' perceptions of the school environment and experiences. *Prevention Science, 15,* 940-950. doi: 10.1007/s11121-013-0432-4

Pas, E. T., Bradshaw, C. P., & Hershfeldt, P. A. (2012). Teacher- and school-level predictors of teacher efficacy and burnout: Identifying potential areas for support. *Journal of School Psychology, 50,* 129-145. doi:10.1016/j.jsp.2011.07.003

Pas, E. T., Cash, A. H., O'Brennan, L., Debnam, K., & Bradshaw, C. P. (2015). Profiles of high school student behavior in the classroom: Associations with teacher behavior management strategies and classroom composition. *Journal of School Psychology, 53,* 137-148. doi: 10.1016/j.jsp.2014.12.005

Peters, C. D., Kranzler, J. H., Algina, J., Smith, S. W., & Daunic, A. P. (2014). Understanding disproportionate representation in special education by examining group differences in behavior ratings. *Psychology in the Schools, 51,* 452-465. doi:10.1002/pits.21761

Pisacreta, J., Tincani, M., Connell, J. E., & Axelrod, S. (2011). Increasing teachers' use of a 1:1 praise-to-behavior correction ratio to decrease student disruption in general education classrooms. *Behavioral Interventions, 26,* 243-260. doi: 10.1002/bin.341

Reglin, G., Akpo-Sanni, J., & Losike-Sedimo, N. (2009). The effect of a professional development classroom management model on at-risk elementary students' misbehaviors. *Education, 133,* 3-18.

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education, 74*(6), 625-637. doi:10.1002/sce.3730740605

Rusby, J. C., Taylor, T. K., & Milchak, C. (2001). *Assessing school settings: Interactions of students and teachers*. Eugene, OR: Oregon Research Institute.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research, 99,* 323-337. doi:10.3200/joer.99.6.323-338

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31,* 351-380. doi:10.1353/etc.0.0007

Siwatu, K. O. (2007). Preservice teachers' culturally responsive teaching self-efficacy and outcome expectancy beliefs. *Teaching and Teacher Education, 23,* 1086-1101. doi:10.1016/j.tate.2006.07.011

Siwatu, K.O. (2009). Student teachers' self-efficacy beliefs regarding culturally responsive teaching and their professed classroom practices. *Teacher Education and Practice, 22,* 323-333.

Siwatu, K.O., & Starker, T. V. (2010). Predicting preservice teachers' self-efficacy to resolve a cultural conflict involving an African American student. *Multicultural Perspectives, 12,* 10-17. doi:10.1080/15210961003641302

Siwatu, K. O. (2011). Preservice teachers' culturally responsive teaching self-efficacy-forming experiences: A mixed methods study. *Journal of Educational Research, 104,* 360-369. doi:10.1080/00220671.2010.487081

Skiba, R. J., Poloni-Staudinger, L., Gallini, S., Simmons, A. B., & Feggins-Azziz, R. (2006). Disparate access: The disproportionality of African American students with disabilities

across educational environments. *Exceptional Children, 72* (4), 411-424. doi:10.1177/001440290607200402

Soodak, L. C., & Podell, D. M. (1993). Teacher efficacy and student problem as factors in special education referral. *Journal of Special Education, 27*(1), 66-81. doi:10.1177/002246699302700105

Soodak, L. C., Podell, D. M., & Lehman, L. R. (1998). Teacher, student, and school attributes as predictors of teachers' responses to inclusion. *Journal of Special Education, 31*, 480-497. doi:10.1177/002246699803100405

Sorrells, A., Schaller, J., & Yang, N. (2004). Teacher efficacy ratings by African American and European American preservice teachers at a historically Black university. *Urban Education, 39*, 509-536. doi:10.1177/0042085904266917

Stapleton, L. M., McNeish, D. M., & Yang, J. S. (2016). Multilevel and single-level models for measured and latent variables when data are clustered. *Educational Psychologist, 51*(3-4), 317-330. doi:10.1080/00461520.2016.1207178

Starko, A. J., & Schack, G. D. (1989). Perceived need, teacher efficacy, and teaching strategies for the gifted and talented. *Gifted Child Quarterly, 33*(3), 118-122. doi:10.1177/001698628903300305

Spanierman, L. B., Oh, E., Heppner, P. P., Neville, H. A., Mobley, M., Wright, C. V., & Navarro, R. (2011). The Multicultural Teaching Competency Scale: Development and initial validation. *Urban Education, 46*, 440-464. doi:10.1177/0042085910377442

Strahan, R., & Gerbasi, K. (1972). Short, homogeneous versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 28*, 191-193. doi:10.1037/t42769-000

Sugai, G., & Horner, R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy, 24,* 23-50. doi:10.1300/j019v24n01_03

Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*, 245-259.

Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8,* 2-8. doi:10.1177/106342660000800101

Sutherland, K. S., Wehby, J. H., & Yoder, P. J. (2002). Examination of the relationship between teacher praise and opportunities for students with EBD to respond to academic requests. *Journal of Emotional and Behavioral Disorders, 10,* 5-13. doi:10.1177/106342660201000102

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. Boston, MA: Allyn and Bacon.

Teo, T. (2010). Structural equation modelling: A primer for music education researchers. *Music Education Research, 12,* 105-118. doi:10.1080/14613800903568270

Tucker, C. M., Porter, T., Reinke, W. M., Herman, K. C., Ivery, P. D., Mack, C. E., & Jackson, E. S. (2005). Promoting teacher efficacy for working with culturally diverse students. *Preventing School Failure, 50,* 29-34. doi:10.3200/psfl.50.1.29-34

Villegas, A. M., & Lucas, T. (2002). Preparing culturally responsive teachers rethinking the curriculum. *Journal of Teacher Education, 53,* 20-32. doi:10.1177/0022487102053001003

Vincent, C. G., Sprague, J. R., & Tobin, T. J. (2012). Exclusionary discipline practices across students' racial/ethnic backgrounds and disability status: Findings from the Pacific Northwest. *Education and Treatment of Children, 35*(4), 585-601. doi:10.1353/etc.2012.0025

Vincent, C. G., Swain-Bradway, J., Tobin, T. J., & May, S. (2011). Disciplinary referrals for culturally and linguistically diverse students with and without disabilities: Patterns resulting from school-wide positive behavior support. *Exceptionality,* 19, 175-190. doi:10.1080/09362835.2011.579936

Wallace Jr, J. M., Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, ethnic, and gender differences in school discipline among US high school students: 1991-2005. *The Negro educational review, 59*(1-2), 47.

Weaver. Gary R. (1986). Understanding and coping with cross-cultural adjustment Stress. In R.M. Paige (Ed). *Cross-cultural orientation. New conceptualizations and applications*. Lanham, MD: University Press of America.

Woolfolk, A., Rosoff, B., & Hoy, W. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and Teacher Education, 6,* 137–148. doi:10.1016/0742-051x(90)90031-y

Worthington, R. L., Mobley, M., Franks, R. P., & Tan, J. (2000). Multicultural counseling competencies: Verbal content, counselor attributions, and social desirability. *Journal of Counseling Psychology, 47,* 460-468. doi:10.1037/0022-0167.47.4.460

Author Biographical Statements

Kristine E. Larson, Ed.D., is a postdoctoral research fellow at the Johns Hopkins University, Bloomberg School of Public Health. Her research focuses on factors related to teachers' ability to manage student behaviors through the use of general and culturally responsive classroom management strategies, and more broadly on improving school climate. She also works as a systems coach to assist principals in using data to improve school climate as part of the Maryland Safe and Supportive Schools (MDS3) research grant, facilities School-Wide Positive Behavioral Interventions and Supports (SW-PBIS) implementation, and uses the Classroom Check-Up (CCU) coaching model to improve teacher practice.

Elise T. Pas, Ph.D., is an Associate Scientist in the Department of Mental Health at the Johns Hopkins University, Bloomberg School of Public Health. Her research examines the effectiveness of evidence- based interventions delivered through a multi-tiered system of supports, coaching interventions, the influence of teacher and school contextual factors on implementation and student outcomes, and the adoption and implementation of positive behavior supports in schools. She holds a doctorate in school psychology from the University of Maryland, College Park.

Catherine P. Bradshaw, Ph.D., M.Ed., is a professor and the Associate Dean for Research and Faculty Development at the Curry School of Education at the University of Virginia. Her primary research interests focus on the development of aggressive behavior and school-based prevention. She has led a number of federally-funded randomized trials of school-based prevention programs, including Positive Behavioral Interventions and Supports (PBIS) and social-emotional learning curricula.

Michael S. Rosenberg is Dean of the School of Education and Professor of Special Education at the State University of New York (SUNY) at New Paltz. He is also Professor Emeritus at Johns Hopkins University, where prior to his current position he was a Professor and department chair in the Department of Special Education and Associate Dean for Research in the School of Education.

Norma L. Day-Vines is a Professor of Counseling and Human Services at the Johns Hopkins University and an affiliate faculty member in the Center for Africana Studies. She has published widely on counseling strategies for working more effectively with culturally and linguistically diverse children and adolescents, with a special emphasis on African American youth. Most notably, Dr. Day-Vines developed the Continuum of Broaching Behavior, a conceptual framework for examining how counselors explore the contextual dimensions of race, ethnicity, and culture.