Hammill Institute on Disabilities

Behavioral Disorders
2023, Vol. 48(2) 97–105

DOI: 10.1177/01987429211067217
journals.sagepub.com/home/bhd
SAGE

# Do Teachers' Perceptions of High Cultural Responsiveness Predict Better Behavioral Outcomes for Students?

**Lindsay M. Fallon, PhD, BCBA-D[1], Margarida B. Veiga, LICSW[1], Annisha Susilo, MEd[1], and Stephen P. Kilgus, PhD[2]**

**Abstract**

Teachers' perceptions of high cultural responsiveness in the classroom may be related to positive behavioral outcomes (e.g., higher academic engagement, lower social risk), but little research has explored this possibility. This article addresses this research gap by building upon findings from a preliminary paper in which these relationships were evidenced. Specifically, we present two interrelated follow-up studies. Study 1 examined the relationship between ratings of 20 U.S. teachers on a measure of cultural responsiveness, the Double Check Self-Refection Tool, and students' observed classroom behavior. Results from multilevel modeling indicated that higher Double Check scores significantly predicted higher academic engagement and lower disruptive behavior for 454 students observed. Study 2 investigated the relationship between ratings of 30 U.S. teachers on the Double Check Self-Refection Tool and ratings of 622 students' risk on the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS). Results indicated higher Double Check scores were associated with lower ratings of students' social and emotional risk. Findings also indicated identification as a Black student and a student with a disability predicted teachers' perceptions of higher risk, consistent with previous research. As results remain preliminary, implications include recommendations for additional research and high-quality professional development to promote teachers' cultural responsiveness.

**Keywords**

cultural responsiveness, behavior, observation, universal screening, teacher training

There has been extensive research surrounding disproportionate rates of racially and ethnically minoritized (REM) students, specifically Black youth, being identified for special education services under the emotional disturbance category of the Individuals with Disabilities Education Improvement Act of 2004 (e.g., Bal et al., 2019). Researchers have reported that teams often place Black students in overly restrictive educational settings with low expectations for learning (Connor, 2017; Ford, 2012; Skiba et al., 2006) which students report as being harmful and stigmatizing (e.g., Banks, 2017). Placement in overly restrictive special education settings is particularly problematic as determination decisions for emotional disturbance can be subjective, lack reliability, and not take into consideration how students' behavior is shaped by educator practices and the school environment (see Sullivan, 2017).

Data traditionally used to support referral to special education for behavioral concerns may reflect interpersonal and systemic racism (e.g., office discipline referrals [ODRs]; Anyon et al., 2018). For instance, as a result of their analysis of nearly half a million ODRs from more than 1,600 schools nationwide, Smolkowski and colleagues found that in comparison with their White peers, Black students were more likely to be referred for discipline (a) for subjectively defined behavior, (b) by their classroom teachers (versus by staff in other parts of the building), and (c) within the first 90 min of the school day. As a result, researchers have encouraged less biased, more proactive data collection efforts (such as universal screening) to be incorporated in a multitiered system of support (MTSS) framework (Raines et al., 2012). However, results from recent studies indicate that teachers may continue to overidentify Black students for being at risk behaviorally on universal screeners, as well (Izumi, 2020).

To address this, there is a need to examine teachers' perceptions of their own antibias and antiracist beliefs as well as their culturally responsive actions (Arneback & Jämte, 2021;

[1]University of Massachusetts Boston, USA
[2]University of Wisconsin-Madison, USA

**Corresponding Author:**
Lindsay M. Fallon, Department of Counseling and School Psychology, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, MA 02125, USA.
Email: lindsay.fallon@umb.edu

Chen et al., 2009). In the two studies herein, we examine the relationship between teachers' perceptions of their cultural responsiveness, their perception of students' risk, and students' observed classroom behavior. We considered the exploration of these relationships integral to strengthening the rationale for targeted teacher professional development aimed at providing culturally responsive practice (Gay, 2018) to ultimately promote better outcomes for youth.

Various teacher assessments of cultural responsiveness exist, including the Assessment of Culturally and Contextually Relevant Supports (ACCReS; Fallon, Cathcart, et al., 2021). The ACCReS is a 35-item scale in which teachers rate their (a) use of equitable classroom practices (ECP subscale; $\omega = .87$), (b) consideration of culture and context in the classroom (CCC subscale; $\omega = .77$), and (c) access to information and support (e.g., data to drive decision-making, relevant professional development; AIS subscale; $\omega = .86$). Items on the CCC subscale primarily target teachers' thinking and beliefs about culturally responsive practice. Items on the AIS subscale are meant to reflect the system-level supports necessary to promote implementing items on the practices described in items on the ECP subscale more successfully. Scores across the three subscales are aggregated to produce a total score, with a higher score indicating greater agreement with items indicating culturally responsive practice.

In Fallon, Veiga, et al. (2021), data from the ACCReS were used in two related studies. In the first study, teachers ($n = 20$) completed the ACCReS and consented to their students' classroom behavior being observed 3 times by university researchers. In the second study, teachers ($n = 30$) completed both the ACCReS and a universal screener, the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS; Kilgus & von der Embse, 2014). Results from multilevel analyses revealed higher ACCReS scores (a) predicted lower levels of observed classwide disruptive behavior (but not academically engaged behavior) in Study 1, and (b) lower risk on the Social Behavior subscale of the SAEBRS ($p = .016$; but not Academic or Emotional Behavior subscales or Total Behavior) in Study 2. Conclusions called for additional research to explore the possibility that teachers who perceive themselves to engage in more culturally responsive practice might perceive students to be at less social risk and see less disruptive behavior in the classroom.

In this article, we replicate the analyses conducted in Fallon, Veiga, et al. (2021) with the Double Check Self-Refection Tool (Hershfeldt et al., 2009; Cronbach's $\alpha = .65$, hereafter referred to as Double Check) in place of the ACCReS. We sought to determine whether the same relationships were observed and perhaps strengthened by a measure that asked teachers about their classroom practice more extensively. The Double Check is a 26-item tool that targets culturally responsive teaching as it pertains to (a) reflective thinking about student's group membership (e.g., "I make tangible efforts [reading, home visits, interviews, student inventories] to "reach out" and understand differences"), (b) development of authentic relationships with students (e.g., "I take genuine interest in the activities and personal lives of others"), (c) effective communication (e.g., "I consistently communicate high expectations"), (d) connecting curriculum content to student culture ("I highlight cultural differences positively during instruction"), and (e) sensitivity to student's cultural and situational messages (e.g., "I am aware of how situations influence behavior [e.g., health, poverty, dress, neighborhood expectations]"). The Double Check is intended to offer teachers the opportunity to reflect critically on their own instructional practice, interactions with students, and personal reactions to students' behaviors (Hershfeldt et al., 2009).

As in Fallon, Veiga, et al. (2021), this brief report presents two related studies with the following research questions:

**Research Question 1 (RQ1):** What is the relationship between teachers' self-assessment ratings of cultural responsiveness (using the Double Check) and classwide observations of student academic engagement and disruptive behavior?

**Research Question 2 (RQ2):** What is the relationship between self-assessment ratings of cultural responsiveness (using the Double Check) and teachers' perceptions of students' academic, emotional, social, and overall risk on a universal screening (specifically the SAEBRS)?

## General Method

### Overview

Below we present two related studies to answer our RQ1 and RQ2.

### Measures

*Double-Check self-assessment.* In both studies, teachers completed the Double Check by responding to each item using a 0 to 4 Likert-type scale: *Regularly in my class and school* (4), *Most of the time in my class or school* (3), *Rarely in my class or school* (2), *Never in my class and school* (1), and *This does not apply to my class and school* (0). Higher aggregate self-ratings on the Double Check indicate higher reported engagement in culturally responsive practices. Internal consistency of the Double Check calculated with responses from the participant sample was relatively high in both Study 1 ($\alpha = .82$) in Study 2 ($\alpha = .87$).

*Demographic questionnaire.* Teachers completed demographic forms to gather information including teachers'

gender, race, years of professional experience, grade level taught. They also reported students' race, English Learner (EL) status, and disability status.

## Study 1

### Participants

Teachers ($n$ = 20) collectively instructing 454 students were recruited from eight Title I schools in two urban public school districts in the U.S. Northeast in Study 1. Teacher and student demographic variables are reported in Table 1. The majority of teachers were female (90%) and White (85%), with 11 years or more teaching experience (65%), most often working in elementary schools (60%). The majority of students were identified by their teachers as REM youth (73%), specifically Latinx (29%), Black (22%), or multiple races/other (23%). The remaining students were identified as White (27%). Twenty-nine percent of students were identified as EL, and 17% were identified with disabilities. All students within both districts received free breakfast and lunch due to their status as high-need under the Healthy, Hunger-Free Kids Act of 2010 (Public Law 111–296). Classroom observations occurred in teachers' general education ($n$ = 9) or inclusion ($n$ = 11) classrooms. The average classroom size was 22 students.

After obtaining approval from the <university's> Institutional Review Board, district and school-based administrators agreed to share study information with teachers via email. All classroom teachers were eligible to participate in the study. Recruitment occurred until 20 teachers consented to participate in the study. During recruitment, potential participants were told that participation would entail an initial meeting to discuss the consent process, three classroom observations, and the completion of study forms (e.g., demographic questionnaire). Participants were offered a US$50 gift card for taking part in the study. Data collection took place in the Winter and Spring of 2019.

### Measures

As described above, teacher participants completed the Double Check as well as a teacher and student demographic form. Two doctoral-level school psychology students also conducted three 30-min classroom observations with each teacher using systematic direct observation (SDO).

*Systematic direct observation.* Two graduate research assistants were trained to conduct observations by the first author using direct training (direct instruction, modeling, and practice with corrective feedback) with videos of students in a classroom. Raters were trained to assess students' classwide academic engagement and disruptive behavior in the videos until 90% overall agreement with the first author was calculated in training sessions. During the study, one graduate research assistant served as the primary observer for each participant's class, and a second graduate research assistant was also present for one of the three observations so that interobserver agreement (IOA) could be calculated. Across participants, IOA was high for academic engagement (97.4%) and disruptive behavior (92.1%) across all observation sessions.

Academic engagement was measured using a 15-s momentary time sampling procedure (Briesch et al., 2015). Observers used an individual-fixed observation scheme, whereby students were observed in the order in which they were seated (Student 1 at Table 1, Student 2 at Table 1, etc.) Academic engagement was defined as any instances in which a student was actively or passively attending to academic instruction or activities (Fallon et al., 2019). This included writing, raising a hand, answering a question, talking about a lesson, listening to the teacher, asking relevant questions, taking notes appropriately, looking at instructional materials, and participating in assigned tasks.

Disruptive behavior was measured using a partial-interval observation scheme whereby observers recorded if a student being observed engaged in disruptive behavior at any point during a 15-s interval. Disruptive behavior was defined as any instance in which a student engaged in behavior that disrupted the classroom functioning or made it difficult for others to learn (e.g., calling out, leaving seat without permission during whole group instruction, talking to peer(s); Fallon et al., 2019).

### Procedures

The first author solicited teachers' participation by obtaining approval from district superintendents and then school principals from two racially and ethnically diverse, large suburban school districts in the northeastern United States. Upon principal approval, classroom teachers were notified about the study. The first author scheduled individual meetings with interested teachers to discuss the study, obtain informed consent, and schedule classroom observations. Once informed consent was secured for teacher participants, parents in participating classrooms were notified about the study and given the option to opt their child out of study participation, however, no parents selected this option.

Data collection involved three classroom observations by a primary observer (graduate research assistant). At the end of the first observation, teachers were provided with a study packet that included teacher and student demographic forms. On the last day of observations, teachers were provided with the Double Check to complete and return to researchers. Then, in addition to being sent a US$50 gift card, participants received a one-page report summarizing the results of observations in their classroom. The first

**Table 1.** Teacher and Student Demographic Data for Participants in Study 1 and Study 2.

| | Demographic characteristics | | | |
|---|---|---|---|---|
| | Study 1 | | Study 2 | |
| Participants | % | *n* | % | *n* |
| Teacher participants | | | | |
| Gender | | | | |
| Female | 90.0 | 18 | 86.7 | 26 |
| Male | 10.0 | 2 | 13.3 | 4 |
| Race and Ethnicity | | | | |
| Asian | 5.0 | 1 | 3.3 | 1 |
| Black or African American | 0.0 | 0 | 3.3 | 1 |
| Latinx | 5.0 | 1 | 0 | 0 |
| Multiracial | 5.0 | 1 | 6.7 | 2 |
| White | 85.0 | 17 | 86.7 | 26 |
| Highest degree earned | | | | |
| Bachelor's | 10.0 | 2 | 6.7 | 2 |
| Masters | 70.0 | 14 | 56.7 | 17 |
| Masters + | 20.0 | 4 | 36.7 | 11 |
| Certification type | | | | |
| General education certification | 75.0 | 15 | 66.7 | 20 |
| Special education certification | 5.0 | 1 | 6.7 | 2 |
| Both | 20.0 | 4 | 26.7 | 8 |
| Years of teaching experience | | | | |
| <1 year | 0 | 0 | 0 | 0 |
| 1–5 years | 5.0 | 1 | 3.3 | 1 |
| 6–10 years | 30.0 | 6 | 40.0 | 12 |
| ≥11 years | 65.0 | 13 | 56.7 | 17 |
| Current grade taught | | | | |
| Elementary (K–5th grades) | 60.0 | 12 | 40.0 | 12 |
| Secondary (6th–8th grades) | 20.0 | 4 | 50.0 | 15 |
| High school (9th–12th grades) | 20.0 | 4 | 10.0 | 3 |
| Student participants | | | | |
| Identified with disability | 17.0 | 77 | 26.2 | 163 |
| English learner | 29.1 | 132 | 25.6 | 159 |
| Race/Ethnicity | | | | |
| Black | 21.8 | 99 | 28.0 | 174 |
| Latinx | 28.6 | 130 | 22.3 | 139 |
| Other or multiple races/ ethnicities | 22.9 | 104 | 23.3 | 145 |
| White | 26.7 | 121 | 26.4 | 164 |

*Note.* These demographics are also presented in Fallon, Veiga, et al. (2021) to answer distinct but related research question.

author offered the opportunity for teachers to debrief and discuss study findings with participants.

## *Statistical Analysis*

To account for the hierarchical nature of the data, a multilevel model analysis was conducted to examine the predictive relationship between teachers' self-rating on the Double Check and student engagement and disruptive behavior in the classroom (Twisk, 2006). The two-level models nested observations (Level 1) within classrooms (Level 2); all models were fit using the lme4 package in R version 3.5.1 (Bates et al., 2015). Due to the small sample size, all models were fit using random intercept only. Although the primary predictor of interest was teacher's Double Check score, the models controlled for other variables that may have influenced student behavior: observation number, teacher's years of experience, school level (elementary, secondary), percentage of EL students, percentage of students identified with disabilities, and percentage of REM students in the class.

To assess the appropriateness of adding Double Check score to the model, we first ran models predicting classwide

**Table 2.** Multilevel Regression Analysis Predicting Classwide Engagement and Disruptive Behavior in Study 1.

| | Classwide academic engagement | | | Classwide disruptive behavior | | |
|---|---|---|---|---|---|---|
| Effects | Coefficient | *SE* | *p* | Coefficient | *SE* | *p* |
| Fixed effects | | | | | | |
| Level 2 variables | | | | | | |
| *Double Check* score | 0.008 | 0.003 | .003** | –0.007 | 0.002 | <.001*** |
| Years of teaching experience | 0.004 | 0.002 | .038* | –0.003 | 0.001 | .020* |
| School level | | | | | | |
| Secondary | –0.021 | 0.053 | .692 | 0.064 | 0.029 | .025* |
| Percentage of EL students | 0.087 | 0.058 | .135 | 0.016 | 0.031 | .614 |
| Percentage of students with disabilities | 0.063 | 0.067 | .344 | –0.045 | 0.036 | .208 |
| Percentage of REM[a] students | –0.146 | 0.106 | .167 | 0.115 | 0.057 | .044* |
| Level 1 variables | | | | | | |
| Observation number | | | | | | |
| Time 2 | 0.008 | 0.020 | .699 | –0.028 | 0.016 | .071 |
| Time 3 | 0.014 | 0.020 | .499 | –0.018 | 0.016 | .24 |
| Random effects | | | | | | |
| Teacher | 0.002 | | | 0.000 | | |
| Residuals | 0.004 | | | 0.002 | | |
| ICC: teacher | 0.33 | | | 0.07 | | |

*Note.* Sixty observations of 20 classrooms across two districts. EL = English Learner; ICC = intraclass correlation coefficients.
[a]REM students indicate students identified as Black, Latinx, other, and/or multiple races/ethnicities.
*$p < .05$. **$p < .01$. ***$p < .001$.

engagement and disruptive behavior without including the Double Check score. Then the Double Check score was added to both models. Both the Akaike information criterion (AIC) and Bayesian information criteria (BIC) were calculated to compare the null models to the models with the Double Check score. Finally, to evaluate the percentage of variance attributable to clusters at each level, we computed intraclass correlation coefficients (ICCs).

## Results

The average score on the Double Check was 90.65 ($SD$ = 6.28, range = 77–102). Results from the analysis of variance revealed that the fit of the model predicting student engagement including the Double Check score (AIC = −114.31, BIC = −91.842) was significantly better than the model predicting student engagement without Double Check score (AIC = −109.21, BIC = −88.778, $p = .007$). Similar results were found for the model predicting classwide disruptive behavior: The model with Double Check score (AIC = −157.86, BIC = −135.39) was a significantly better fit than the null model (AIC = −145.79, BIC = −125.36, $p = .0001$).

Teachers' Double Check score was found to be a significant predictor of both classwide academic engagement and disruptive behavior. Higher Double Check self-ratings predicted significantly higher classwide academic engagement ($p = .003$). Teachers' years of experience was also a significant predictor of classwide academic engagement ($p = .038$), such that more experience predicted higher classwide engagement. Higher Double Check scores also predicted lower classwide disruptive behavior ($p < .001$). Percentage of REM students ($p = .008$), teacher experience ($p = .020$), and school level ($p = .010$) were also significant predictors of lower classwide disruptive behavior. Both the values for ICC for student academic engagement (ICC = 0.33) and student disruptive behavior (ICC = 0.07) were above .05, which was determined to be substantial and indicate that the shared variance at Level 2 was significant. Detailed results of the regression analyses are presented in Table 2.

## Study 2

### Participants

A total of 30 teachers and 622 students from a total of nine schools participated in the second study. To be included in Study 2, teachers were required to have completed the SAEBRS for each student in the class observed. Twelve teachers from Study 1 participated and an additional 18 teachers were recruited for Study 2 (see "Procedures" section). Most participants were female (86.7%), White (86.7%), and had 11 or more years of experience (56.7%). The majority of students were identified as REM youth (74%), including Black (28%) and Latinx (22%). The remainder were identified as White. About one quarter of students were identified as EL (26%) and/or with a disability (26%). As in Study 1, all students within both districts

received free breakfast and lunch due to their status as high-need under the Healthy, Hunger-Free Kids Act of 2010 (Public Law 111–296). Complete demographic characteristics of teachers and students can be found in Table 1.

### Measures

Teachers completed the SAEBRS (Kilgus & von der Embse, 2014), a brief universal screener designed to be used in Grades K–12 to assess student's functioning in three subscales: Social Behavior (six items), Academic Behavior (six items), and Emotional Behavior (seven items). There is also a Total Behavior score on the SAEBRS (reflecting all 19 items), which is an estimate of overall functioning. The items reflect behaviors found to be highly correlated with social and academic success (Eklund et al., 2017). The tool is efficient, requiring approximately 1–3 min to screen each student. Teachers completed the scale using a categorical 4-point scale from 0 (*never*) to 3 (*almost always*) on the degree to which an item was true for the child being screened. Previous research supports the internal consistency of the SAEBRS ($\alpha$ = .79–.94; Kilgus & von der Embse, 2014).

### Procedures

The additional sample of teachers ($n$ = 18) was recruited using the same procedures outlined in Study 1. Teachers who indicated interest in participation first consulted with the first author. Once informed consent was secured, participants were asked to complete the Double Check, SAEBRS, and teacher and student demographic forms.

### Data Analysis

Multilevel modeling was used to assess the relationship between teachers' Double Check score and perceptions of student risk on the SAEBRS. Teachers' scores for individual students (Level 1) were nested within classrooms (Level 2); all analyses were again conducted in R Version 3.5.1 using the lme4 package (Bates et al., 2015). Covariates included in the models were similar to those in Study 1: teacher's years of experience, school level, and student demographic information (i.e., race, EL status, disability status). Similarly, null models without Double Check score were run first to assess the appropriateness of including Double Check score in models predicting teacher's ratings of student risk on each of the four subscales of the SAEBRS. We also calculated AIC, BIC and ICCs.

### Results

The average score on the Double Check scale was 91.19 ($SD$ = 7.02, range = 75–102). Although Double Check score did not improve the fit of the model predicting teacher's perceptions of academic risk ($p$ = .06), disability status was found to be a significant predictor of academic risk ($p < .001$). Also, Double Check score did not significantly predict perceptions of risk on the Academic Behavior subscale.

Double Check score significantly improved the fit of models predicting emotional ($p$ = .007), social risk ($p <$ .001), and overall ($p$ = .019) risk. Higher self-ratings on the Double Check predicted lower perceptions of student risk on the Social Behavior ($p < .001$) and Emotional Behavior ($p$ = .002) SAEBRS subscales. Double Check score also significantly predicted lower SAEBRS Total Behavior ($p$ = .014) score. Students in secondary school were more likely to be identified as at risk on the Emotional Behavior subscale ($p < .044$). Being identified as a Black student was also a significant predictor of teachers' perceptions of higher risk in the Social Behavior ($p$ = .004) subscale. Finally, students identified as having a disability were significantly more likely to be considered more at risk across SAEBRS ($p < .001$) subscales. Detailed results of the analyses are summarized in Table 3. The ICC values for Academic Behavior (ICC = 0.06), Emotional Behavior (ICC= 0.19), Social Behavior (ICC =0.06), and Total Behavior (ICC = 0.10) indicated substantial shared variance at Level 2.

## General Discussion

Overall, results indicated teachers who perceived themselves to engage in more culturally responsive practice also had better classwide behavioral outcomes and perceived students to be at less social-emotional risk.

In Study 1, higher Double Check scores (indicating higher teacher perceptions of their culturally responsive classroom practice) predicted both higher academic engagement and lower classwide disruptive behavior. This extended the findings from Fallon, Veiga, et al. (2021), which noted a relationship between teachers' perception of culturally responsive practice as measured by the ACCReS and lower disruptive behavior. It may be that there is more academic engagement and less disruption in classrooms in which teachers who are reflective in their thinking about student's group membership, seek to develop authentic relationships with students, work to communicate effectively, connect the curriculum content to student culture, and are sensitive to cultural and situational messages (e.g., "I am aware of how situations influence behavior").

In Study 2, higher Double Check ratings were associated with lower perceptions of student risk as indicated by the Social Behavior, Emotional Behavior, and Total Behavior scale, but not associated with lower risk on the Academic Behavior subscale of the SAEBRS. In other words, teachers who perceived themselves as more culturally responsive perceived students to have fewer challenges with social behavior (e.g., disruptive behavior, arguing, outbursts) and

**Table 3.** Multilevel Regression Analysis Predicting SAEBRS Subscales in Study 2.

| | SAEBRS subscale | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Social behavior | | | Academic behavior | | | Emotional behavior | | | Total behavior scale | | |
| Effects | Coefficient | *SE* | *p* | Coefficient | *SE* | *p* | Coefficient | *SE* | *p* | Coefficient | *SE* | *p* |
| Fixed effects | | | | | | | | | | | | |
| Level 2 variables | | | | | | | | | | | | |
| *Double Check* score | 0.138 | 0.036 | <.001*** | 0.070 | 0.037 | .060 | 0.130 | 0.046 | .005** | 0.269 | 0.110 | .014* |
| Teaching experience | 0.012 | 0.041 | .766 | –0.026 | 0.043 | .539 | –0.034 | 0.053 | .520 | –0.093 | 0.127 | .462 |
| School level | | | | | | | | | | | | |
| Secondary | –0.176 | 0.540 | .745 | 0.000 | 0.559 | 1.000 | –1.402 | 0.695 | .044** | –1.347 | 1.662 | .417 |
| Level 1 variables | | | | | | | | | | | | |
| Student race | | | | | | | | | | | | |
| Black | –1.422 | 0.494 | .004** | –0.180 | 0.509 | .724 | 0.558 | 0.429 | .194 | 0.131 | 1.313 | .921 |
| Latinx | –0.888 | 0.532 | .095 | –0.465 | 0.548 | 0395 | 0.433 | 0.455 | .341 | –0.502 | 1.403 | .720 |
| Other[a] | –0.937 | 0.520 | .072 | 0.189 | 0.535 | .725 | 0.070 | 0.443 | .875 | 0.490 | 1.370 | .721 |
| Student EL status | 0.034 | 0.469 | .942 | –0.102 | 0.483 | .833 | 0.312 | 0.423 | .461 | 0.118 | 1.268 | .926 |
| Student disability status | –1.838 | 0.439 | <.001*** | –2.024 | 0.453 | <.001*** | –1.840 | 0.401 | <.001*** | –6.036 | 1.193 | <.001*** |
| Random effects | | | | | | | | | | | | |
| Teacher | 1.08 | | | 1.18 | | | 2.68 | | | 12.71 | | |
| Residuals | 16.41 | | | 17.36 | | | 11.54 | | | 112.02 | | |
| ICC: teacher | 0.06 | | | 0.06 | | | 0.19 | | | 0.10 | | |

*Note.* Ratings for 622 students from 30 classrooms across two districts. SAEBRS = Social, Academic, and Emotional Behavior Risk Screener; ICC = intraclass correlation coefficients; EL = English learner.

[a]Other includes identified as a race/ethnicity of than Black, Latinx, White, or includes students identified as multiple races/ethnicities.

*$p < .05$. **$p < .01$. ***$p < .001$.

emotional behavior (e.g., withdrawal, sadness, fearfulness) among their students, but not necessarily less academic risk (e.g., distractedness). Considering the findings from Study 1, it may be that teachers with higher Double Check scores may have reported lower levels of risk on the SAEBRS as their students were typically more engaged and less disruptive, yet additional research is needed to explore this possibility. Additional research might also target why teachers rated Black students as being at increased Social Behavioral risk. This finding is aligned with previous research (e.g., Izumi, 2020), but warrants empirical exploration as it seems possible that racism and bias may continue to play a role in teachers' perceptions of students' social risk, even for teachers who consider themselves to implement high levels of culturally responsive practice in the classroom.

### *Limitations*

Although these studies provide preliminary evidence of a relationship between teachers' perceptions of their cultural responsiveness and student outcomes, several limitations should be considered when interpreting results. In Study 1, SDOs of students' classwide behavior occurred over the course of three 30-min observations. Data collected during these observations may not have been representative of typical classroom behavior due to the students' and/or teachers' reactivity to the presence of the observers as well as the limited number of observations. Future research might include more observations conducted randomly throughout the day to ensure that data collected are representative of the students' typical behavior. Second, research related to the psychometric properties of the Double Check is limited. However, we administered the Double Check as it has been used in subsequent studies exploring the association between observed and self-reported culturally proficient teaching practices (e.g., Debnam et al., 2015). We calculated the internal consistency and found it to be relatively high for the two samples in this report. Yet there remains a need for additional research exploring the technical adequacy of the Double Check for broader use in research and practice. Furthermore, although teachers were encouraged to respond honestly to the Double Check items, social desirability bias may have affected teachers' responses. Due to the nonevaluative nature of the study and the range of responses on items on the measure, concern about social desirability bias is limited. Future studies might include a social desirability scale to assess respondents' concerns with social approval. Relatedly, it is also very possible that because these studies relied on teachers' opinion of their practices and students' social, academic, and emotional risk, they may misperceive their own actions or their students' behavior. Simply, there is a risk that teachers will not see limitations to their own practice and instead associate risk or "misbehavior" with students (i.e., a fundamental attribution error). Future studies might incorporate observations of teachers' practices to evaluate if observers' ratings on the Double Check items are accurate as well as ratings of students' behavior and teachers' perceptions of student risk (i.e., specific to each student within the context of their SAEBRS scores). In addition, a future study might include responses from students pertaining to how they might rate their teacher on measures of cultural responsiveness (e.g., ACCReS, Double Check), or interview students about their perceptions of teacher practices and how they have experienced their teachers to be in interpersonal interactions. Including data on student perception has not, to our knowledge, been done frequently in quantitative analyses of teachers' culturally responsive practice but is warranted. Finally, researchers conducting future studies should collect additional demographic data (e.g., students' socioeconomic status, specific disability classification or disorders such as behavior disorders) to include as variables in analyses, as well as include a larger sample size representing schools outside of the northeast United States to generalize findings.

### *Implications*

Implications for the current findings include both opportunity for additional research and application for practice. Researchers should continue to address the influence of bias on data collected and interpreted in schools. This may be particularly relevant in the context of decision-making about eligibility for special education services for emotional disturbance given extensive evidence in the literature regarding disproportionality by student race (Bal et al., 2019).

Researchers might also continue to explore whether teachers might consider their efforts to be antibiased and culturally responsive (Chen et al., 2009) as a critical step in promoting positive behavioral outcomes for students. In schools, data from teacher self-assessments such as the Double Check might be used to inform targeted professional development to support teachers to strengthen their practice. For instance, if teachers rate items related to developing authentic relationships with students lower, coaching teachers to engage in positive interactions with students that promote trust may ultimately support student behavior and impact how teachers perceive students' behavioral risk. Additional research is needed to further explore the preliminary findings presented in this brief report. This scholarship is urgently needed to promote equitable, supportive environments for all youth, specifically REM students.

### *Conclusion*

In this article, we present preliminary evidence of a relationship between teachers' perceptions of their own cultural responsiveness, perception of students' risk, and students' observed classroom behavior. Teachers who perceive themselves to be more culturally responsive might design educational environments in which youth are seen as being at less

risk and in which students are more actively engaged in classroom instruction. Results justify the need for teacher professional development aimed at promoting culturally responsive practice.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The U.S. Department of Education's Institute of Education Sciences supported this research through Grant R324B170010 to the University of Massachusetts Boston. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

## References

Anyon, Y., Lechuga, C., Ortega, D., Downing, B., Greer, E., & Simmons, J. (2018). An exploration of the relationships between student racial background and the school sub-contexts of office discipline referrals: A critical race theory analysis. *Race Ethnicity and Education*, *21*(3), 390–406. https://doi.org/10.1080/13613324.2017.1328594

Arneback, E., & Jämte, J. (2021). How to counteract racism in education–A typology of teachers' anti-racist actions. *Race Ethnicity and Education*. Advance online publication. https://doi.org/10.1080/13613324.2021.1890566

Bal, A., Betters-Bubon, J., & Fish, R. E. (2019). A multilevel analysis of statewide disproportionality in exclusionary discipline and the identification of emotional disturbance. *Education and Urban Society*, *51*(2), 247–268. https://doi.org/10.1177/0013124517716260

Banks, J. (2017). "These people are never going to stop labeling me": Educational experiences of African American male students labeled with learning disabilities. *Equity & Excellence in Education*, *50*(1), 96–107. https://doi.org/10.1080/10665684.2016.1250235

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Briesch, A. M., Hemphill, E. M., Volpe, R. J., & Daniels, B. (2015). An evaluation of observational methods for measuring response to classwide intervention. *School Psychology Quarterly*, *30*(1), 37–49. https://doi.org/10.1037/spq0000065

Chen, D. W., Nimmo, J., & Fraser, H. (2009). Becoming a culturally responsive early childhood educator: A tool to support reflection by teachers embarking on the anti-bias journey. *Multicultural Perspectives*, *11*(2), 101–106. https://doi.org/10.1080/15210960903028784

Connor, D. J. (2017). Who is responsible for the racialized practices evident within (special) education and what can be done to change them? *Theory Into Practice*, *56*(3), 226–233. https://doi.org/10.1080/00405841.2017.1336034

Debnam, K. J., Pas, E. T., Bottiani, J., Cash, A. H., & Bradshaw, C. P. (2015). An examination of the association between observed and self-reported culturally proficient teaching practices. *Psychology in the Schools*, *52*(6), 533–548. https://doi.org/10.1002/pits.21845

Eklund, K., Kilgus, S., von der Embse, N., Beardmore, M., & Tanner, N. (2017). Use of universal screening scores to predict distal academic and behavioral outcomes: A multilevel approach. *Psychological Assessment*, *29*(5), 486–499. https://doi.org/10.1037/pas0000355

Fallon, L. M., Cathcart, S. C., Johnson, A. H., Minami, T., O'Keeffe, B. V., DeFouw, E. R., & Sugai, G. (2021). *A self-assessment of culturally relevant practice to inform educators' professional development*. Manuscript under review.

Fallon, L. M., Collier-Meek, M. A., & Kurtz, K. D. (2019). Feasible coaching supports to promote teachers' classroom management in high-need settings. *School Psychology Review*, *48*(1), 3–17. http://doi.org/10.17105/SPR-2017-0135.V48-1

Fallon, L. M., Veiga, M., Susilo, A., Robinson-Link, P., Berkman, T., Minami, T., & Kilgus, S. P. (2021). Exploring the relationship between teachers' perceptions of culturally relevant supports, student risk, and classroom behavior. *Psychology in the Schools*. Advance online publication. https://doi.org/10.1002/pits.22568

Ford, D. Y. (2012). Culturally different students in special education: Looking backward to move forward. *Exceptional Children*, *78*(4), 391–405. https://doi.org/10.1177/001440291207800401

Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.

Hershfeldt, P. A., Sechrest, R., Pell, K. L., Rosenberg, M. S., Bradshaw, C. P., & Leaf, P. J. (2009). Double-check: A framework of cultural responsiveness applied to classroom behavior. *TEACHING Exceptional Children Plus*, *6*(2), 2–18.

Izumi, J. (2020). *Detecting and explaining differential item functioning on the social, academic, and emotional behavior risk screener* (Doctoral Dissertation). University of Missouri.

Kilgus, S. P., & von der Embse, N. P. (2014). *Social, academic, and emotional behavior risk screener (SAEBRS)*. Theodore J. Christ & Colleagues.

Raines, T. C., Dever, B. V., Kamphaus, R. W., & Roach, A. T. (2012). Universal screening for behavioral and emotional risk: A promising method for reducing disproportionate placement in special education. *The Journal of Negro Education*, *81*(3), 283–296. https://doi.org/10.7709/jnegroeducation.81.3.0283

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. https://doi.org/10.1177/001440290607200402

Smolkowski, K., Girvan, E. J., McIntosh, K., Nese, R. N., & Horner, R. H. (2016). Vulnerable decision points for disproportionate office discipline referrals: Comparisons of discipline for African American and White elementary school students. *Behavioral Disorders*, *41*(4), 178–195. https://doi.org/10.17988/bedi-41-04-178-195.1

Sullivan, A. L. (2017). Wading through quicksand: Making sense of minority disproportionality in identification of emotional disturbance. *Behavioral Disorders*, *43*(1), 244–252. https://doi.org/10.1177/0198742917732360

Twisk, J. W. (2006). *Applied multilevel analysis: A practical guide for medical researchers*. Cambridge University Press.