Article

Hammill Institute on Disabilities

Behavioral Disorders
2021, Vol. 46(2) 96–107

DOI: 10.1177/0198742920911651
journals.sagepub.com/home/bhd
SAGE

Classroom Behavior of Students With or At Risk of EBD: Student Gender Affects Teacher Ratings But Not Direct Observations

Amanda W. Sheaffer, PhD[1], Caitlyn E. Majeika, PhD[2], Allison F. Gilmour, PhD[3], and Joseph H. Wehby, PhD[1]

Abstract

As the field moves toward adaptive and individualized behavior intervention, it is important to identify and consider relevant student characteristics as potential levers (i.e., critical factors) for improving intervention effectiveness. Motivated by previous findings suggesting that behavioral profiles and teachers' perceptions of students with problem behavior vary by student gender, we evaluated gender differences in teacher ratings and direct observations of classroom problem behavior for elementary students with or at risk of emotional and behavioral disorder (EBD; $N = 352$). We found significant gender differences for teacher ratings of problem behavior, social skills, and academic competence. However, we did not find significant differences between male and female students on direct observation measures of behavior or reading skills. These findings provide evidence suggesting that student gender may be a critical factor in addressing classroom problem behavior. Future work in this area will help guide researchers and practitioners in considering how student gender and related teachers' perceptions may inform intervention selection and implementation for students with or at risk of EBD.

Keywords

perceptions, teacher(s), disorders, behavior(s), elementary, age group

Students with emotional and behavioral disorder (EBD) share common characteristics including the frequent display of problem behavior, academic difficulty, and social skills deficits. Although intervention development and selection for students with or at risk of EBD has historically targeted specific skill deficits, scholars have also recently focused on identifying key correlates, referred to as critical factors, that may be important considerations in the intervention planning process (Majeika et al., 2020; Sterrett et al., 2020; Wehby & Kern, 2014). Importantly, past research finds that students with or at risk of EBD may exhibit different characteristics depending on their gender (e.g., Docherty et al., 2016), making gender a potentially important critical factor to consider during intervention planning. In the following sections, we provide a brief overview of students with EBD and present a rationale for examining gender of students with or at risk of EBD as a potential critical factor to inform future intervention development, before presenting the research questions driving the present study.

Characteristics of Students With or At Risk of EBD

By definition, students with or at risk of EBD are prone to convergent behavioral, academic, and social risk factors and exhibit maladaptive behavior that prevents them from forming appropriate relationships and inhibits their learning. Examples include high rates of disruptive classroom behavior (Erickson et al., 2006; Kelly & Shogren, 2014; Turtura et al., 2014), poor social skills (Gresham et al., 2011; McEvoy & Welker, 2000; Quinn et al., 1999), and academic skill deficits (Bub et al., 2007; Gage et al., 2017;

[1]Vanderbilt University, Nashville, TN, USA
[2]University of North Texas, Denton, USA
[3]Temple University, Philadelphia, PA, USA

Corresponding Author:
Amanda W. Sheaffer, Department of Special Education, Vanderbilt University, OMC, 110 Magnolia Circle, Nashville, TN 37203, USA.
Email: amanda.w.sheaffer@vanderbilt.edu

Lane et al., 2008). Students with or at risk of EBD are more likely than typically developing peers to have low rates of on-task behavior (Erickson et al., 2006; Kelly & Shogren, 2014) and high rates of avoidance of academic tasks (Erickson et al., 2006; Turtura et al., 2014). Due to the bidirectional association between academics and behavior, students with or at risk of EBD are often characterized by low academic achievement in reading and math with little to no growth over time (Bub et al., 2007; Gage et al., 2017; Lane et al., 2008). In addition to behavioral and academic deficits, students with or at risk of EBD also have social skills deficits that inhibit positive social dynamics at school (Gresham et al., 2011).

Without timely and effective intervention, students with or at risk of EBD face dismal life outcomes. Rates of problem behavior in the early years often increase with age and, in turn, increase the likelihood for negative long-term outcomes (Henry et al., 2012; Montague et al., 2005; Murray & Farrington, 2010; Nelson et al., 2006). Problem behavior during adolescence is predictive of later antisocial behavior (Henry et al., 2012; Lo et al., 2002), substance abuse (McGue & Iacono, 2005), and delinquency (Sedlak & McPherson, 2010). Moreover, low social competence in childhood is associated with greater internalizing and externalizing problem behaviors during adolescence (Bornstein et al., 2010). In summary, the cumulative literature on students with or at risk of EBD presents a particularly bleak picture and highlights a need for research that focuses on improving outcomes for this population of students.

Given the comorbid academic, social, and behavioral problems, and devastating immediate and life outcomes associated with EBD, researchers often target behavioral, social, and academic outcomes when developing and evaluating interventions for students with or at risk of EBD (e.g., Gage et al., 2012). Despite the focus on these highly salient outcome variables, students may benefit from interventions that are adapted based on other potentially relevant variables, also known as critical factors. The identification of critical factors may represent a fruitful approach to inform decisions about behavioral interventions and improve intervention effectiveness for this population.

Critical Factors for Students With or At Risk of EBD

A critical factor is a variable that is likely to impact an individual's response to intervention and may be identified at the student or teacher level (Lei et al., 2012; Wehby & Kern, 2014). Student-level critical factors include skill deficits, baseline rates of behavior, and function of behavior. For example, a student who engages in attention-maintained problem behavior is more likely to benefit from mentoring or a check-in check-out intervention that increases access to positive adult feedback than a student who engages in escape-maintained problem behavior (Fairbanks et al., 2007). Critical factors also exist at the teacher level and may include teacher training, teacher buy-in, and teacher's perception (Majeika et al., 2020). For example, a teacher who is well trained and wants to implement an intervention may be more likely to do so with fidelity than a teacher who has a negative attitude about the intervention (Miramontes et al., 2011).

Knowledge of critical factors can be used at key decision points during intervention implementation to (a) match a student to an intervention and (b) adjust intervention components to better meet the needs of a student. Thus, the use of critical factors during intervention planning and implementation has the potential to increase the efficiency and effectiveness of behavioral interventions. As reviewed in the following section, research suggests that student gender may be another example of a salient critical factor for students with or at risk of EBD. Determining whether and how gender is associated with outcomes of students with or at risk of EBD could inform the development of gender-based interventions that more precisely meet student needs.

Student Gender as a Critical Factor

Gender-Based Developmental Trajectories of Problem Behavior

In the literature on the trajectory of problem behavior for students with or at risk of EBD, females have been historically understudied. Although decades of work evaluating the developmental trajectories of problem behavior were largely evaluated on male populations (e.g., Patterson & Yoerger, 1993), these models have been broadly applied across both males and females (Frances et al., 1994). However, researchers have more recently suggested that males and females with or at risk for EBD demonstrate different developmental patterns of problem behavior (Fontaine et al., 2009; Silverthorn & Frick, 1999). For example, a greater proportion of girls initiate problem behavior during adolescence than during childhood, and these girls experience a unique pathway characterized by lifelong difficulties (Bardone et al., 1998; Fontaine et al., 2009; Moffitt et al., 2002; Odgers et al., 2008; Silverthorn & Frick, 1999). This adolescent-onset pathway for girls indicating persistent issues into adulthood stands in contrast to the trajectory applied to adolescent-onset males who tend to discontinue problem behavior after adolescence (Moffitt, 1993; Silverthorn & Frick, 1999). In addition to highlighting developmental differences by gender, researchers have also argued that males and females learn problem behavior through differential gender roles and processes (Hagan et al., 1987; Heimer, 1996; Heimer & Coster, 1999; Weerman et al., 2016). Collectively, the literature underscores gender

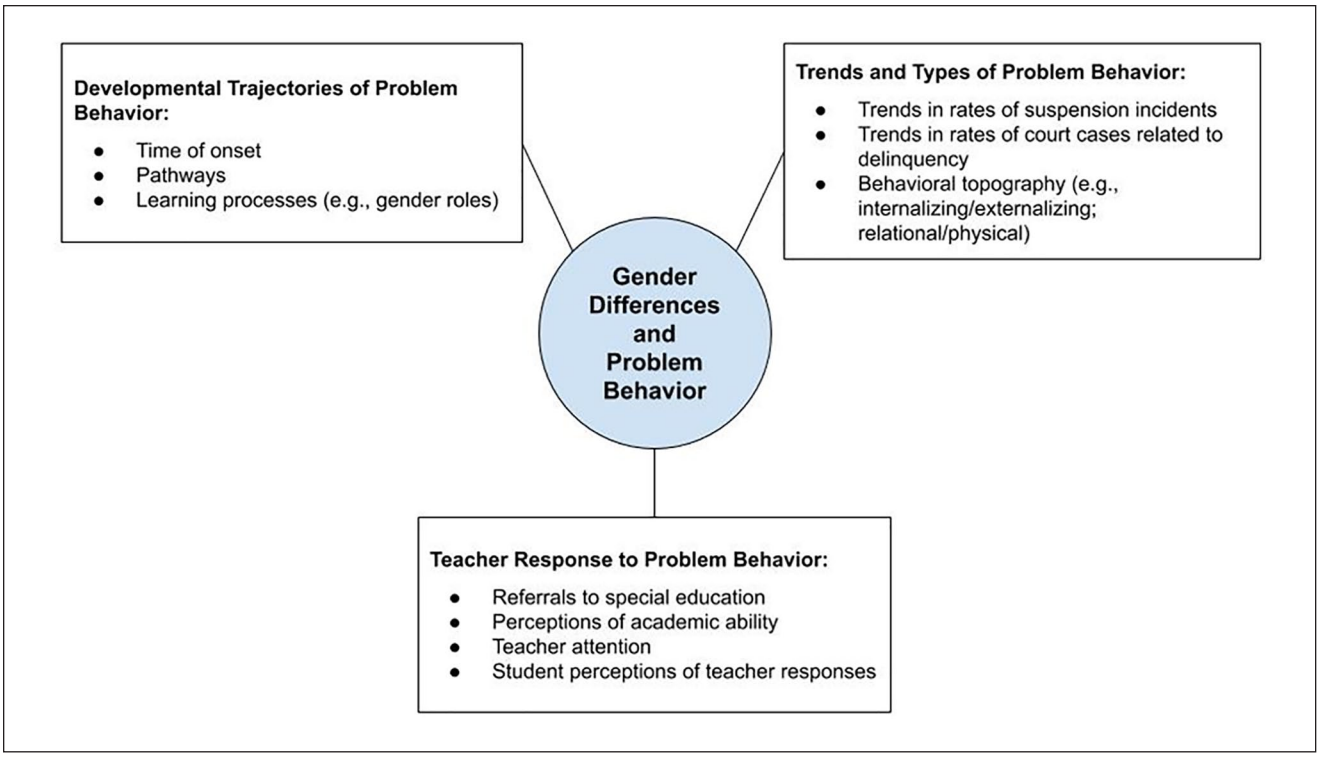

Figure 1. Gender differences and problem behavior.

as a potentially important factor in the development of problem behavior (see Figure 1).

Changing Trends and Types of Problem Behavior for Females

Although national discipline data show that female students exhibit less problem behavior at school than males, females represent hundreds of thousands of school-based infractions each year (U.S. Department of Education, 2017). The Civil Rights Project of University of California Los Angeles reported that suspension rates are rising for female populations at a rate higher than males (Losen & Skiba, 2010). Moreover, the National Center for Juvenile Justice reported an increase in the proportion of court cases representing females over the past 30 years (Sickmund et al., 2018).

Findings in the literature also suggest that problem behavior may be descriptively different by gender. Prior research shows that females are more likely than males to present internalizing behaviors, such as depression or anxiety, and display antisocial behavior that is often motivated by interpersonal conflict (Docherty et al., 2016; Ehrensaft, 2005). In one study on patterns of aggression, researchers found females to be more likely to display relational aggression than males, who are more likely to display overt, physical aggression (Crick & Grotpeter, 1995). These findings are supported by Card and colleagues (2008) who conducted a meta-analysis on gender differences for direct (physical and verbal) and indirect (relational, covert, social) aggression. Findings indicated boys were more likely than girls to engage in direct aggression; whereas other research has found that girls are more likely than boys to engage in indirect aggression (Archer & Coyne, 2005; Crapanzano et al., 2010). Similarly, in a qualitative study of middle school special education teachers of students with EBD, teachers reported that female students presented more covert problem behaviors stemming from personal grudges than their male counterparts (Rice et al., 2008). These findings show that relational aggression may be a gender-specific factor driving problem behavior in females and suggest that researchers and practitioners may need to consider distinctive, gender-based topographies of problem behavior when selecting and implementing behavioral interventions.

Gender and Teachers' Perception

Just as student-level characteristics and behavioral trends show notable gender differences, teacher-level factors may also vary by student gender. Specifically, researchers have found that teachers' perceptions of and interactions with students may differ according to gender (see Beaman et al., 2006). For example, in a study of 202 kindergarten teachers, teachers gave attention to unsociable behaviors when demonstrated by boys more often than when demonstrated by

girls (Arbeau & Coplan, 2007). Students' perceptions of teacher behavior reinforce these findings. For example, boys report that teachers are more lenient toward girls and report more reprimands toward boys than girls in class (Consuegra, 2015; Younger et al., 1999).

Patterns in special education referrals provide further evidence of varying teachers' perceptions by gender. Teachers may refer male students to special education more frequently than females despite similar rates of teacher-rated academic and behavioral deficits (Henning-Stout, 1993; Skårbrevik, 2002; Vogel, 1990). These findings suggest that girls have to demonstrate more severe deficits than males before being referred for special education or other support services.

Relatedly, teachers' perceptions of academic ability or achievement may also vary by gender. Decades of research has indicated teachers hold gendered stereotypes of students' academic proficiency in core content areas (Tiedemann, 2002). Regardless of student performance on academic assessments, teacher ratings of student competence in mathematics and reading vary by student gender (Robinson & Lubienski, 2011; Robinson-Cimpian et al., 2014). Moreover, extensive research has been done to explore how these gendered views affect teacher practices (Fennema, 1990; Li, 1999; Robinson-Cimpian et al., 2014; Tiedemann, 2002). However, findings are inconsistent and vary by factors such as student age, prior academic achievement, student behavior, and teacher attitudes (Fennema et al., 1990; Helwig et al., 2001; Robinson-Cimpian et al., 2014). The inconclusive findings related to gender-based teacher perceptions of academic competence bolster the need for continued consideration of student gender as a relevant variable to study.

Purpose

To date, the cumulative literature indicates differential (a) trajectories and topographies of problem behavior, (b) patterns of teacher behavior, and (c) teachers' perceptions of student behavior and academic ability by gender. Across studies, authors used a variety of measures; however, few studies analyzed findings in a way that could potentially also offer insight into the relation between teachers' perceptions and observed behaviors. Considering the literature indicating student gender as an important student- and teacher-level factor related to problem behavior (e.g., teachers' perception by student gender) and the potential need for tailored interventions, further examination of gender-based, teacher-rated behavioral profiles and directly measured behaviors of students with or at risk of EBD is warranted to help inform the development of gender-based interventions for students with or at risk of EBD.

In this study, we explored differences between male and female students exhibiting persistent classroom problem behavior. To address limitations in the existing literature in which many studies include only single measures of classroom problem behavior, we included data from both teacher ratings and direct observations. Specifically, we evaluated gender differences in (a) teacher ratings of social skills, problem behavior, internalizing behaviors, and academics and (b) direct observations of aggression, negative talk, and disengagement, and a direct assessment of students' reading achievement.

We asked the following research questions for a sample of elementary students with or at risk of EBD in urban school districts:

Research Question 1: Do teachers' perceptions of behavioral characteristics and academic ability differ by gender for students exhibiting persistent problem behavior?

Research Question 2: Do direct observations of classroom behaviors differ by gender for students exhibiting persistent problem behavior?

Method

Sample

We used data collected from a multisite study evaluating the impact of a teacher-level classroom management intervention in urban elementary schools in Tennessee, Minnesota, and Virginia (Maggin et al., 2011). The 2-year study included two cohorts of students with or at risk of EBD. Participants were identified from 30 schools distributed across three large metropolitan school districts from three regions of the United States that expressed interest in participation. Schools were identified by participating districts based on school demographics, special education enrollment, and frequency of office discipline referrals. Given site differences in the service-delivery models, the sample was stratified by district and matched pairs of schools were developed within each district based on the presence of a special education classroom and the grade levels served by interested participants. Within pairs of schools, random assignment was made to either an intervention or comparison group (see Maggin et al., 2011, for additional details about the sample; the results of this study are not reported here). Districts ranged in size from 23,200 to 70,140 with average racial demographics of 24% White, 58% Black, 12% Hispanic, 4% Asian, and 2% American Indian; 12% limited English proficiency; 15% received special education services; and 69% of students qualified for free or reduced lunch.

Following approval from the Institutional Review Board and participating school districts, consent forms were sent home to all students in the classrooms of teachers participating in the classroom management study. For those

Table 1. Sample Demographics.

	Full sample (n = 352)		Female (n = 80)		Male (n = 272)	
Variable	Statistic	Range	Statistic	Range	Statistic	Range
Average age	7.91 (1.49)	5.07–11.16	7.80 (1.42)	5.09–10.98	7.94 (1.51)	5.17–11.16
% Black	75.57		75.37		76.25	
% Other minority	0.57		1.25		0.37	
% FRL	90.34		86.25		91.54	
% SPED	73.01		66.25		75.00	
% ELL	1.47		2.70		1.13	
Negative talk	0.06 (0.12)	0–1.17	0.04 (0.04)	0–0.27	0.07 (0.13)	0–1.17
Aggression	0.01 (0.02)	0–0.15	0.01 (0.02)	0–0.1	0.01 (0.03)	0–0.15
Disengaged	13.91 (13.44)	0–78.31	11.72 (11.14)	0–57.47	14.56 (14.00)	0–78.31
Social skills	85.65 (14.32)	38–130	80.99 (13.57)	52–111	87.04 (14.27)	38–130
Problem behavior	116.97 (13.96)	76–145	123.25 (13.70)	85–144	115.09 (13.50)	76–145
Academic skills	90.06 (14.25)	62–133	86.32 (12.71)	65–116	91.19 (14.51)	62–133
WJ-LWID	87.41 (16.25)	47–130	88.38 (15.66)	47–130	87.14 (16.43)	48–93
Internalizing behaviors	62.21 (10.08)	37–94	61.71 (10.57)	37–94	62.36 (9.95)	38–89

Note. Standard deviations are in parentheses. Negative talk and aggression are average rates. Disengagement is the average percentage of the observation. Social skills, problem behavior, and academic skills are the standard scores from teacher ratings on the SSRS. FRL = free/reduced lunch; SPED = receiving special education services; ELL = English language learner; WJ-LWID = Woodcock Johnson Letter-Word Identification standard score; SSRS = Social Skills Rating System.

students for whom approval was obtained, screening into the project occurred in two ways. In general education classrooms, at-risk status was determined by the Systematic Screening for Behavior Disorders (SSBD; Walker et al., 1990). Inclusion criteria for general education students were based on recommendations provided in the SSBD manual. General education students were identified as at risk if they were reported to have either (a) five or more Critical Events or (b) at least one Critical Event, an Adaptive Behavior score of at least 30, and a Maladaptive Behavior score of at least 35. Students enrolled in special education, self-contained classrooms for students with EBD were automatically included in the study based on their assignment to these classrooms specifically for behavioral issues.

The data set started with 419 students, but 67 students were missing at least one demographic variable. After removing these students, we were left with a sample of 352 students. We included students who had at least one dependent variable: 331 students had teacher ratings of social skills, 343 students had teacher ratings of problem behavior, 340 students had teacher ratings of academic skills, and 347 students had data for the direct observation measures. Of the entire sample, 22.7% were female, which closely aligns with national statistics indicating roughly 27.5% of school-aged students with EBD are female (U.S. Department of Education, 2018). The students in our sample were taught by 115 teachers in 30 different schools. Each teacher had 1–16 students (M = 3.06, SD = 2.64); each school had 1–10 teachers (M = 3.83, SD = 2.49).

The first column of Table 1 presents participant demographic characteristics. Participants were enrolled in kindergarten through fourth grades, reflecting the grade span targeted for the original intervention study. The majority of the sample was Black, qualified for free or reduced lunch, and was classified by the school as receiving special education services. Very few students were English language learners or from other minority groups. The characteristics of students' teachers aligned with national estimates from the National Center for Education Statistics (NCES): the sample was majority female (sample = 91%, nationally = 89%), primarily fully licensed (sample = 89%, nationally = 91%), and had an average of 13 years of experience (nationally = 14 years of experience).

Measures

Rating of student behavior. The *Social Skills Rating System* (SSRS; Gresham & Elliott, 1990) is an assessment used to evaluate social and behavioral characteristics in students from the perspective of multiple raters (i.e., parents, teachers, students) at the preschool, elementary, and secondary levels. The assessment consists of three rating scales: Social Skills, Academic Competence, and Problem Behavior. For purposes of this study, we focused solely on the teacher version of the rating scale. The internal consistencies of the problem behavior, social skills, and academic competence subscales of the teacher rating form ranged from .88 to .95 (Gresham et al., 2011). The measure also has demonstrated adequate construct and criterion-related validity (Elliott et al., 1988).

We also used data collected from the *Teacher's Report Form Internalizing Scale* (TRF; Achenbach & Resorla, 2001) to evaluate students' internalizing behaviors. This measure asks teachers to rate students on items addressing social withdrawal, somatic complaints, and anxiety/depression. This measure has shown acceptable test–retest reliability, internal consistency, and convergent validity (Achenbach & Resorla, 2001). Teachers completed the rating forms after school had been in session for more than 2 months and within 8 weeks of a student being enrolled in the study.

Student reading achievement. We included in our analyses a direct assessment of students' reading skills to complement the teachers' ratings of Academic Competence. In the original study, researchers used the Woodcock Johnson III Letter-Word Identification subtest to assess students' early reading skills (Woodcock et al., 2001). The test manual reports reliability estimates of .92–.93 for this subtest.

Direct observation of student behavior. Research assistants (RAs) directly observed problem behavior for each student up to four times ($M = 3.92$, $SD = 0.45$, range = 1–4) using the Multiple Option Observation System for Experimental Studies (MOOSES) program (Tapp & Wehby, 2000) during 15-min observation sessions. The MOOSES program is a handheld, computer-based observation system for collection of discrete events and durational measures in real time. Each observation took place at the time in the school day during English language arts or math instruction within a 2-week period, with efforts to complete all observations during the same week. We collected the observation data in December, January, or February. The direct observation variables we focused on were negative talk, disengagement, and aggression. Negative talk and aggression were measured by frequency counts. Disengagement was measured as the percentage of observed time that the student was disengaged. We define these behaviors below.

Negative talk. Negative talk was defined as statements or vocalizations made with the intent to provoke, annoy, pester, mock, whine, complain, tattle, or make fun of another. This category also included threats of physical aggression against a person or property, arguing or disagreeing with another person (as in protest), as well as any verbal refusal to comply with a command from the teacher.

Disengagement. Disengagement was defined as instances when a student is not participating in an approved/assigned activity. This category included looking around the room, out of seat/walking around during instructional activity, disrupting others, talking to peers when he or she is not supposed to, and sleeping.

Aggression. Aggression was defined as deliberate physical contact that is potentially harmful to self, others, or property, and posturing or a gesturing that is intended to provoke another.

Training and interobserver agreement (IOA). To ensure reliability and consistency between observers during data collection, IOA was established during training and live coding sessions. Prior to live observations, each RA was trained on the MOOSES program using practice videos. To be considered reliable on the practice videos, the RAs have to meet a criterion of at least 80% agreement with the master video files across three consecutive sessions. We used the MOOSES program to calculate IOA by dividing the number of agreements by the sum of agreements and disagreements. For frequency counts (i.e., negative talk, aggression), observers agreed when indicating the onset of the behavior within a 3-s window of each other. For measures of duration (i.e., disengagement), observers had to record the start and end of the behavior within 3 s of each other to be in agreement. Once RAs met the training criterion, they engaged in practice sessions in classroom settings. Pairs of RAs coded until meeting a reliability criterion of 80% with an expert coder across three consecutive sessions. Once meeting reliability criteria for the practice videos and classrooms, RAs were able to code student data in the target classrooms. Because this was a multi-year study, these training procedures were in place for all RAs at the start of every year.

IOA was measured during 20% of sessions across each participant. During these sessions, a second observer was present to record observational data. Agreement estimates were calculated using the traditional agreements/(agreements + disagreements) formula. The average agreement across behaviors was 93.4% with a range of 87% to 100%.

Data Analysis

We first examined correlations between the dependent variables to ensure that each measure captured different constructs (see Table 2). There were statistically significant, moderate correlations between negative talk, aggression, and disengagement. Similarly, teacher ratings of student social skills, academic skills, and problem behaviors were moderately correlated. Teacher ratings of internalizing behaviors had low to moderate correlations with negative talk, problem behavior, and social skills. Aggression was not correlated with any teacher ratings. Disengagement was positively correlated with problem behavior and negatively correlated with both social skills and academics. Negative talk was also positively correlated with teacher ratings of problem behavior. Reading achievement and teachers' ratings of students' academic skills were positively correlated. The low to moderate correlations among the dependent

Table 2. Correlations Between Outcomes.

Variable	Negative talk	Aggression	Log disengagement	Problem behavior	Social skills	Academics	WJ-LWID	Internalizing behavior
Negative talk	1.00							
Aggression	.50***	1.00						
Log disengagement	.23***	.30***	1.00					
Problem behavior	.11*	.06	.22***	1.00				
Social skills	−.02	.02	−.27***	−.55***	1.00			
Academics	.06	.03	−.19**	−.28***	.59***	1.00		
WJ-LWID	−.03	−.03	−.02	.00	−.06	.33***	1.00	
Internalizing	.12*	.03	.02	.20***	−.13*	−.01	−.11*	1.00

Note. Average rate of negative talk, average rate of aggression, and average percentage of time disengaged are used in these correlations.
WJ-LWID = Woodcock Johnson Letter-Word Identification standard score.
*$p < .05$. **$p < .01$. ***$p < .001$.

variables supported that each addressed a different construct, and that fitting separate models for each dependent variable was appropriate.

We then examined the distribution of each outcome. Aggression and negative talk were both zero-inflated. We used multilevel Poisson regression when we conducted the analyses for these two outcomes because these dependent variables did not meet assumptions of normality. Disengagement was skewed right. We added a constant of 1 and took the log of disengagement to better approximate a normal distribution. Teacher ratings of social skills, academic skills, problem behavior, and internalizing behavior were all normally distributed, as was student reading achievement. We used linear multilevel models for log disengagement, academic skills, problem behavior, and internalizing behavior.

The students in our sample were nested within teachers ($n = 115$) and schools ($n = 30$). We used multilevel modeling so that the standard errors and significance tests would not be biased due to the nonindependence of students. We determined the number of required levels based on the intraclass correlation (ICC) and included random intercepts when the ICC for the level was above 0. We used students' average rate of negative talk, average rate of aggression, and average log disengagement as our dependent variables instead of modeling multiple observations nested within students. We fit three-level models for two outcomes, average rate of negative talk, and the average percentage of the observation that a student was disengaged (students nested in teachers, and teachers nested in schools). The three-level models included a teacher random intercept and school random intercept. We used a two-level model with students nested in teachers to address rate of aggression because the ICC suggested that variation at the school level was nearly zero. We fit three-level models for the teacher ratings of social skills, problem behavior, academic skills, and internalizing behavior, and students' reading achievement. The three-level models included teacher and school random intercepts.

We fit one model for each dependent variable using the specifications described above. In each model we included an indicator for if the student was female (1 = female, 0 = male), our primary predictor of interest, and we included control variables for student age (continuous variable), race/ethnicity (1 = Black, 0 = White), free/reduced lunch (FRL) eligibility (1 = eligible; 0 = not eligible for FRL), and for receiving special education services as indicated by having an Individualized Education Program (IEP) (1 = has IEP; 0 = does not have IEP). We included these control variables because of the concern that they could bias teacher ratings of students and to account for some demographic differences between the male and female students in our sample. The student characteristics were Level 1 predictors, but were Level 2 predictors in the models using direct observation data. All of the variables were fixed coefficients. We converted the coefficients from the linear models to effect sizes using Cohen's *d* and converted the coefficients from the Poisson models to incidence rate ratios (IRRs) to improve interpretability.

Results

In Table 3, we present the results for models examining teacher ratings of behavior. We identified multiple statistically significant differences between female and male students. Teachers rated female students −5.28 points below male students in social skills on average ($p < .01$), an effect size of −0.37. Teachers gave higher average ratings in problem behavior to females compared with males ($d = .55$; $p < .001$) and lower ratings in academic skills ($d = -.27$; $p < .05$), differences that were statistically significant even after accounting for other student characteristics. Teachers rated female and male students similarly on internalizing behavior ($d = .08$; $p > .05$).

Table 3. Results for Multilevel Models Comparing Female and Male Students on Teacher Rating Outcomes.

Variable	Social skills	Problem behavior	Academic skills	Internalizing behavior
Female	−5.28** (1.71)	7.40*** (1.60)	−3.83* (1.72)	0.76 (1.20)
Age	1.76*** (0.54)	1.03 (0.52)	0.18 (0.55)	1.11* (0.40)
Black	−1.40 (1.78)	2.80 (1.63)	2.24 (1.78)	0.73 (1.23)
FRL	−3.19 (2.50)	−3.05 (2.31)	−7.01** (2.51)	−1.40 (1.70)
SPED	4.48* (1.84)	−7.81*** (1.75)	3.78* (1.85)	0.36 (1.39)
School RE	6.34* (7.41)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Teacher RE	24.56* (11.37)	36.69* (11.83)	33.85* (13.08)	33.41* (8.24)
Residual var.	148.04* (13.11)	129.22* (11.54)	154.50* (15.53)	66.60 (6.07)
n	331	343	340	337

Note. Outcomes are teacher ratings on the SSRS. FRL = free/reduced lunch; SPED = receiving special education services; RE = random effects; SSRS = Social Skills Rating System.
*$p < .05$ or, for RE, 95% confidence interval does not include 0. **$p < .01$. ***$p < .001$ level.

Table 4. Results for Multilevel Models Comparing Female and Male Students on Directly Observed and Assessed Outcomes.

Variable	Negative talk	Aggression	Disengagement	WJ-LWID
Female	−0.59 (0.65)	−0.51 (1.52)	−0.21 (0.12)	−0.58 (1.74)
Age	−0.04 (0.15)	−0.12 (0.38)	0.01 (0.04)	−3.90 (0.55)
Black	0.12 (0.56)	0.65 (1.61)	0.04 (0.12)	0.74 (1.81)
Free/reduced lunch	0.12 (0.85)	0.28 (2.40)	0.03 (0.16)	−5.39 (7.53)
Special education	0.54 (0.59)	0.10 (3.66)	−0.30 (0.14)	−18.45*** (5.07)
School RE	0.00 (0.00)	—	0.03* (0.06)	28.16* (13.57)
Teacher RE	0.00 (0.00)	0.00 (0.00)	0.34* (0.10)	23.18* (13.94)
n	347	347	347	338

Note. Multilevel Poisson regression was used for negative talk and aggression models. Aggression is a two-level model; the model would not converge with the inclusion of the school random effect. Disengagement is log disengagement with a constant of 1 to account for right skew in the distribution. WJ-LWID = Woodcock Johnson Letter-Word Identification standard score; RE = random effects.
*$p < .05$ level or, for RE, the 95% confidence interval does not include 0.

In Table 4, we present the results from the multilevel models we used to compare female and male students on the three directly observed behaviors and reading skills. On average, after accounting for student characteristics, female students had slightly lower average rates of negative talk (IRR = 0.55; $p > .05$) and aggression (IRR = 0.60; $p > .05$), and had lower average log disengagement than their male peers ($d = -.11$; $p > .05$). However, none of these differences were statistically significant. Female students had slightly lower average reading scores than male students in this sample, but this difference was not statistically significant ($d = -.04$; $p > .05$). Overall, we identified no significant differences between female and male students on directly observed or assessed outcomes.

Discussion

Students with or at risk of EBD face some of the most detrimental school- and life-based outcomes of all students with disabilities. To better understand the behavioral profiles of these students, and given the potential relevance of gender as a critical factor, we examined gender-based academic, social, and behavioral characteristics of elementary students with or at risk of EBD. Based on our analyses, gender differences existed on global, teacher-rated measures of social skills, problem behavior, and academic competence; however, direct observations of negative talk, aggression, and engagement indicated no differences by gender.

Teacher Bias

In this study, findings show that girls with or at risk of EBD were rated by their teachers as having poorer social skills, lower academic competence, and worse behavior than their male counterparts. Simply put, teachers viewed males and females with problem behavior differently despite similar classroom behaviors. These findings may provide support for Tiet and colleagues' (2001) theory of a gender paradox in which a smaller number of females with or at risk of EBD represent the most severe infractions. Collectively, the results from our study (a) support the notion that student gender may affect how teachers perceive students with problem behavior and (b) have potential implications related to teacher behavior.

Admittedly, implications for teacher behavior assume a link between teachers' perceptions and behaviors, a relationship moderately supported in the literature. Although the literature in this area is limited, prior work shows that teachers' perceptions are associated with teacher practices. For example, teachers' perceptions have been linked to differential patterns of referral to special education (Morgan et al., 2010; Skårbrevik, 2002; Vogel, 1990), behavioral feedback in the classroom (Beaman et al., 2006), attitudes toward students (Arbeau & Coplan, 2007; Consuegra, 2015), educational decisions (e.g., class placement; Zucker & Prieto, 1977), expectations for student achievement (McKown & Weinstein, 2008), and implementation of inclusive practices (De Boer et al., 2010).

Given the association between teachers' perception and teacher behavior, our findings of differential ratings may suggest that teachers respond to problem behavior differently by student gender, perhaps due to specific assumptions teachers have about these behaviors and gender norms. If teacher response to student misbehavior varies systematically by student gender, intervention and discipline procedures may be enacted inequitably and leave a particular gender group without appropriate behavioral support. Although we found no meaningful differences in observed behavior between males and females with and at risk of EBD, teachers rated females as exhibiting greater problem behavior and lower social and academic skills. Our findings add to previous evidence showing that teachers attend to males with problem behavior more often than females (Arbeau & Coplan, 2007), suggesting that gender bias may be a related factor for students with or at risk of EBD. If teachers view females with or at risk of EBD more harshly than similar males and therefore do not provide the intervention needed to improve outcomes for these females, then girls with or at risk of EBD may receive disproportionately fewer supports than males. Following this logic, our findings indicate a need for training or intervention at the teacher level to account for gender bias when working with students with or at risk of EBD.

Limitations and Future Directions

The findings of this study should be interpreted in light of a few key limitations. To start, the majority of our sample was male (77%), Black (75%), and receiving special education services (73%). The limited variation in type of participant inhibits the generalizability of our findings. Teacher biases based on gender, race, or disability may have affected the teacher ratings, though we attempted to account for this possibility through the addition of control variables in our statistical models. Our sample was also limited by the percentage of female students (22.7%), though the percentage of female students in our sample is only slightly lower than 27.5% of school-aged students with EBD who are female (U.S. Department of Education, 2018). Ideally, we would have over-sampled female students to examine gender differences. Our present sample had power over 0.8, the traditionally acceptable level, for differences between genders of over 0.36 *SD*. Additional work needs to be conducted to examine whether patterns of gender differences are present in samples with more females, across a variety of races, and for more students who are at risk of a disability but not formally diagnosed.

In addition, teacher ratings of social skills, problem behavior, and academic competence may reflect different constructs of student behavior than those shown by direct observation measures of negative talk, aggression, and engagement. Gender differences on teacher ratings and direct observation measures may capture different aspects of student behavior. Because the variables in this study were not precisely aligned across direct assessment and teacher perception measures, we cannot ignore that alternate explanations for our findings exist. For example, we cannot rule out the possibility that teachers' perceptions may not have varied if ratings scale items focused specifically on negative talk, aggression, and engagement. As such, findings must be interpreted with caution. In addition, we did not focus extensively on positive behaviors. We captured some positive behaviors with the social skills rating form, but we did not capture positive behaviors with direct observation. Another limitation, due to the measures used, is that the observations and teacher ratings were not completed at the same time. Although our goal in this study was not to assess convergent validity, students' behaviors may have changed in specific ways from the time in which they were rated to when they were observed. However, this would only influence the comparative analyses if these changes happened in systematically different ways across female and male students on the direct observation measures. We recommend future researchers (a) use multiple measures to assess both positive and negative student behavior, (b) carefully align measures with the behavioral constructs associated with the research questions, and (c) attend to the timing of their assessments. A final limitation is the relatively brief length of observation sessions (15 min) and the limited number per participants ($M = 3.92$, $SD = 0.45$, range = 1–4).

Due to these limitations, additional future research is needed to replicate our analyses and provide further evidence to confirm that gender differences in teachers' perceptions exist for students with or at risk of EBD. To aid interpretation of findings, future work should also explore the association between teachers' perception and teacher behavior. Specifically, researchers should execute a multimethod approach by using observation measures and teacher ratings to ensure analyses across multiple aspects of problem behavior, include measures of teacher behavior, and evaluate gender differences on all measures. Work to

this end will facilitate a more comprehensive understanding of the intersection of student gender, teachers' perception, and teacher practice.

Conclusion

As the field moves toward adaptive and individualized behavior interventions, it is important to consider students' characteristics and teachers' perceptions that may inform intervention. Our findings provide evidence suggesting gender may be a critical factor that warrants further study as it relates to problem behavior. Future work in this area will help guide researchers and practitioners in considering how gender may inform intervention selection and implementation. Furthermore, an understanding of gender as a potential critical factor may contribute to the refinement of systematic processes used to better align and tailor behavior interventions for students with or at risk of EBD.

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 author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding for the fourth author was provided, in part, by a grant from the U.S. Department of Education, #H325H14000.

ORCID iD

Caitlyn E. Majeika https://orcid.org/0000-0002-8917-655X

References

Achenbach, T. M., & Resorla, L. (2001). *Manual for the ASEBA school-age forms & profiles: An integrated system of multi-informant assessment*. ASEBA.

Arbeau, K. A., & Coplan, R. J. (2007). Kindergarten teachers' beliefs and responses to hypothetical prosocial, asocial, and antisocial children. *Merrill-Palmer Quarterly*, *53*(2), 291–318.

Archer, J., & Coyne, S. M. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review*, *9*(3), 212–230. https://doi.org/10.1207/s15327957pspr0903_2

Bardone, A. M., Moffitt, T. E., Caspi, A., Dickson, N., Stanton, W. R., & Silva, P. A. (1998). Adult physical health outcomes of adolescent girls with conduct disorder, depression, and anxiety. *Journal of the American Academy of Child & Adolescent Psychiatry*, *37*(6), 594–601. https://doi.org/10.1097/00004583-199806000-00009

Beaman, R., Wheldall, K., & Kemp, C. (2006). Differential teacher attention to boys and girls in the classroom. *Educational Review*, *58*(3), 339–366. https://doi.org/10.1080/00131910600748406

Bornstein, M. H., Hahn, C. S., & Haynes, O. M. (2010). Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: Developmental cascades. *Development and Psychopathology*, *22*(4), 717–735. https://doi.org/10.1017/S0954579410000416

Bub, K. L., McCartney, K., & Willett, J. B. (2007). Behavior problem trajectories and first-grade cognitive ability and achievement skills: A latent growth curve analysis. *Journal of Educational Psychology*, *99*(3), 653–670. https://doi.org/10.1037/0022-0663.99.3.653

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development*, *79*(5), 1185–1229. https://doi.org/10.1111/j.1467-8624.2008.01184.x

Consuegra, E. (2015). Gendered teacher-student classroom interactions in secondary education: Perception, reality and professionalism [Doctoral thesis]. Vrije Universiteit Brussel.

Crapanzano, A. M., Frick, P. J., & Terranova, A. M. (2010). Patterns of physical and relational aggression in a school-based sample of boys and girls. *Journal of Abnormal Child Psychology*, *38*(4), 433–445. https://doi.org/10.1007/s10802-009-9376-3

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development*, *66*(3), 710–722. https://doi.org/10.2307/1131945

De Boer, A., Pijl, S. J., & Minnaert, A. (2010). Attitudes of parents towards inclusive education: A review of the literature. *European Journal of Special Needs Education*, *25*(2), 165–181. https://doi./org/10.1080/08856251003658694

Docherty, M., Boxer, P., Veysey, B. M., & Ostermann, M. (2016). Gender differences in prevalence of internalizing and externalizing symptoms in a justice-referred sample of youth. *Journal of Juvenile Justice*, *5*(2), 104–120.

Ehrensaft, M. K. (2005). Interpersonal relationships and sex differences in the development of conduct problems. *Clinical Child and Family Psychology Review*, *8*(1), 39–63. http://doi.org/10.1007/s10567-005-2341

Elliott, S. N., Gresham, F. M., Freeman, T., & McCloskey, G. (1988). Teacher and observer ratings of children's social skills: Validation of the Social Skills Rating Scales. *Journal of Psychoeducational Assessment*, *6*(2), 152–161. https://doi.org/10.1177/073428298800600206

Erickson, M. J., Stage, S. A., & Nelson, J. R. (2006). Naturalistic study of the behavior of students with EBD referred for functional behavioral assessment. *Journal of Emotional and Behavioral Disorders*, *14*(1), 31–40. https://doi.org/10.1177/10634266060140010301

Fairbanks, S., Sugai, G., Guardino, D., & Lathrop, M. (2007). Response to intervention: Examining classroom behavior support in second grade. *Exceptional Children*, *73*(3), 288–310. https://doi.org/10.1177/001440290707300302

Fennema, E. (1990). Teachers beliefs and gender differences in mathematics. In E. Fennema & G. C. Leder (Eds.), *Mathematics and gender* (pp. 169–187). Teachers' College Press.

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys,

and mathematics. *Educational Studies in Mathematics*, *21*(1), 55–69. https://doi.org/10.1007/BF00311015

Fontaine, N., Carbonneau, R., Vitaro, F., Barker, E. D., & Tremblay, R. E. (2009). Research review: A critical review of studies on the developmental trajectories of antisocial behavior in females. *Journal of Child Psychology and Psychiatry*, *50*(4), 363–385. https://doi.org/10.1111/j.1469-7610.2008.01949

Frances, A., Pincus, H. A., & First, M. B. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*. American Psychiatric Association.

Gage, N. A., Adamson, R., MacSuga-Gage, A. S., & Lewis, T. J. (2017). The relation between the academic achievement of students with emotional and behavioral disorders and teacher characteristics. *Behavioral Disorders*, *43*(1), 213–222. https://doi.org/10.1177/0198742917713211

Gage, N. A., Lewis, T. J., & Stichter, J. P. (2012). Functional behavioral assessment-based interventions for students with or at risk for emotional and/or behavioral disorders in school: A hierarchical linear modeling meta-analysis. *Behavioral Disorders*, *37*(2), 55–77. https://doi.org/10.1177/019874291203700201

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system: Manual*. American Guidance Service.

Gresham, F. M., Elliott, S. N., Vance, M. J., & Cook, C. R. (2011). Comparability of the social skills rating system to the social skills improvement system: Content and psychometric comparisons across elementary and secondary age levels. *School Psychology Quarterly*, *26*(1), 27–44. https://doi.org/10.1037/a0022662

Hagan, J., Simpson, J., & Gillis, A. R. (1987). Class in the household: A power-control theory of gender and delinquency. *American Journal of Sociology*, *92*(4), 788–816. https://doi.org/10.1086/228583

Heimer, K. (1996). Gender, interaction, and delinquency: Testing a theory of differential social control. *Social Psychology Quarterly*, *59*(1), 39–61. https://doi.org/10.2307/2787118

Heimer, K., & Coster, S. D. (1999). The gendering of violent delinquency. *Criminology*, *37*(2), 277–318.

Helwig, R., Anderson, L., & Tindal, G. (2001). Influence of elementary student gender on teachers' perceptions of mathematics achievement. *The Journal of Educational Research*, *95*(2), 93–102. https://doi.org/10.1080/00220670109596577

Henning-Stout, M. (1993). The relationship between teachers' perceptions of girls' and boys' behaviors and referral for special education. *Special Services in the Schools*, *7*(1), 91–105. https://doi.org/10.1300/J008v07n01_06

Henry, K. L., Knight, K. E., & Thornberry, T. P. (2012). School disengagement as a predictor of dropout, delinquency, and problem substance use during adolescence and early adulthood. *Journal of Youth and Adolescence*, *41*(2), 156–166. https://doi.org/10.1007/s10964-011-9665-3

Kelly, J. R., & Shogren, K. A. (2014). The impact of teaching self-determination skills on the on-task and off-task behaviors of students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *22*(1), 27–40. https://doi.org/10.1177/1063426612470515

Lane, K. L., Barton-Arwood, S. M., Nelson, J. R., & Wehby, J. (2008). Academic performance of students with emotional and behavioral disorders served in a self-contained setting. *Journal of Behavioral Education*, *17*(1), 43–62. https://doi.org/10.1007/s10864-007-9050-1

Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D., & Murphy, S. A. (2012). A "SMART" design for building individualized treatment sequences. *Annual Review of Clinical Psychology*, *8*, 21–48. https://doi.org/10.1146/annurev-clinpsy-032511-143152

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research*, *41*(1), 63–76. https://doi.org/10.1080/0013188990410106

Lo, Y. Y., Loe, S. A., & Cartledge, G. (2002). The effects of social skills instruction on the social behaviors of students at risk for emotional or behavioral disorders. *Behavioral Disorders*, *27*(4), 371–385. https://doi.org/10.1177/019874290202700409

Losen, D. J., & Skiba, R. J. (2010). *Suspended education: Urban middle schools in crisis*. The Civil Rights Project.

Maggin, D. M., Wehby, J. H., Partin, T. C., Robertson, R., & Oliver, R. M. (2011). A comparison of the instructional context for students with behavioral issues in self-contained and general education classrooms. *Behavioral Disorders*, *36*, 84–99.

Majeika, C. E., Bruhn, A. L., Sterrett, B. I., & McDaniel, S. A. (2020). Reengineering Tier 2 interventions for responsive decision making: An adaptive intervention process. *Journal of Applied School Psychology*. Advance online publication. https://doi.org/10.1080/15377903.2020.1714855

McEvoy, A., & Welker, R. (2000). Antisocial behavior, academic failure, and school climate: A critical review. *Journal of Emotional and Behavioral Disorders*, *8*(3), 130–140. https://doi.org/10.1177/106342660000800301

McGue, M., & Iacono, W. G. (2005). The association of early adolescent problem behavior with adult psychopathology. *American Journal of Psychiatry*, *162*(6), 1118–1124. https://doi.org/10.1007/s10519-006-9061-z

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, *46*(3), 235–261. https://doi.org/10.1016/j.jsp.2007.05.001

Miramontes, N. Y., Marchant, M., Heath, M. A., & Fischer, L. (2011). Social validity of a positive behavior interventions and support model. *Education and Treatment of Children*, *34*(4), 445–468. https://doi.org/10.1353/etc.2011.0032

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*(4), 674–701. https://doi.org/10.1037/0033-295X.100.4.674

Moffitt, T. E., Caspi, A., Harrington, H., & Milne, B. J. (2002). Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. *Development and Psychopathology*, *14*(1), 179–207.

Montague, M., Enders, C., & Castro, M. (2005). Academic and behavioral outcomes for students at risk for emotional and behavioral disorders. *Behavioral Disorders*, *31*(1), 84–94. https://doi.org/10.1177/019874290503100106

Morgan, P. L., Frisco, M. L., Farkas, G., & Hibel, J. (2010). A propensity score matching analysis of the effects of special education services. *The Journal of Special Education*, *43*(4), 236–254. https://doi.org/10.1177/0022466908323007

Murray, J., & Farrington, D. P. (2010). Risk factors for conduct disorder and delinquency: Key findings from longitudinal

studies. *The Canadian Journal of Psychiatry*, *55*(10), 633–642. https://doi.org/10.1177/070674371005501003

Nelson, J. R., Benner, G. J., Stern, N., & Stage, S. A. (2006). Interrelationships among language skills, externalizing behavior, and academic fluency and their impact on the academic skills of students with ED. *Journal of Emotional and Behavioral Disorders*, *14*(4), 209–216. https://doi.org/10.1177/10634266060140040401

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology*, *20*(2), 673–716. https://doi.org/10.1017/S0954579408000333

Patterson, G. R., & Yoerger, K. (1993). Developmental models for delinquent behavior. In S. Hodgins (Ed.), *Mental disorder and crime* (pp. 140–172). SAGE.

Quinn, M. M., Kavale, K. A., Mathur, S. R., Rutherford, R. B., Jr., & Forness, S. R. (1999). A meta-analysis of social skill interventions for students with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *7*(1), 54–64. https://doi.org/10.1177/106342669900700106

Rice, E. H., Merves, E., & Srsic, A. (2008). Perceptions of gender differences in the expression of emotional and behavioral disabilities. *Education and Treatment of Children*, *31*(4), 549–565. https://doi.org/10.1353/etc.0.0035

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, *48*(2), 268–302. https://doi.org/10.3102/0002831210372249

Robinson-Cimpian, J. P., Lubienski, S. T., Ganley, C. M., & Copur-Gencturk, Y. (2014). Teachers' perceptions of students' mathematics proficiency may exacerbate early gender gaps in achievement. *Developmental Psychology*, *50*(4), 1262–1281. https://doi.org/10.1037/a0035073

Sedlak, A. J., & McPherson, K. (2010). *Survey of youth in residential placement: Youth's needs and services* [SYRP report]. Westat.

Sickmund, M., Sladky, A., & Kang, W. (2018). *Easy access to juvenile court statistics: 1985-2017*. https://www.ojjdp.gov/ojstatbb/ezajcs/

Silverthorn, P., & Frick, P. J. (1999). Developmental pathways to antisocial behavior: The delayed-onset pathway in girls. *Development and Psychopathology*, *11*(1), 101–126. https://doi.org/10.1017/S0954579499001972

Skårbrevik, K. (2002). Gender differences among students found eligible for special education. *European Journal of Special Needs Education*, *17*(2), 97–107. https://doi.org/10.1080/08856250210129038a

Sterrett, B. I., McDaniel, S. A., Majeika, C. E., & Bruhn, A. L. (2020). Using evidence informed strategies to adapt Tier 2 interventions. *Journal of Applied School Psychology*. Advance online publication. https://doi.org/10.1080/15377903.2020.1714856

Tapp, J., & Wehby, J. (2000). Observational software for laptop computers and optical bar code time wands. In T. Thompson, D. Felce, & F. Symons (Eds.), *Behavioral observation: Technology and applications in developmental disabilities* (pp. 71–82). Brookes.

Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics*, *50*(1), 49–62. https://doi.org/10.1023/A:1020518104346

Tiet, Q. Q., Wasserman, G. A., Loeber, R., McReynolds, L. S., & Miller, L. S. (2001). Developmental and sex differences in types of conduct problems. *Journal of Child and Family Studies*, *10*(2), 181–197. https://doi.org/10.1023/A:1016637702525

Turtura, J. E., Anderson, C. M., & Boyd, R. J. (2014). Addressing task avoidance in middle school students: Academic behavior check-in/check-out. *Journal of Positive Behavior Interventions*, *16*(3), 159–167. https://doi.org/10.1177/1098300718778032

U.S. Department of Education. (2017). *Number of students receiving selected disciplinary actions in public elementary and secondary schools, by type of disciplinary action, disability status, sex, and race/ethnicity: 2013-14* (Digest of Educational Statistics Table 233.27). https://nces.ed.gov/programs/digest/d17/tables/dt17_233.27.asp

U.S. Department of Education. (2018). *Children 3 to 21 years old served under Individuals with Disabilities Education Act (IDEA), Part B, by age group and sex, race/ethnicity, and type of disability: 2017-18* (Digest of Education Statistics Table 204.50). https://nces.ed.gov/programs/digest/d18/tables/dt18_204.50.asp

Vogel, S. A. (1990). Gender differences in intelligence, language, visual-motor abilities, and academic achievement in students with learning disabilities: A review of the literature. *Journal of Learning Disabilities*, *23*(1), 44–51. https://doi.org/10.1177/002221949002300111

Walker, H. M., Severson, H., & Feil, E. G. (1990). *Systematic screening for behavior disorders (SSBD)*. Sopris West.

Weerman, F. M., Bernasco, W., Bruinsma, G. J., & Pauwels, L. J. (2016). Gender differences in delinquency and situational action theory: A partial test. *Justice Quarterly*, *33*(7), 1182–1209. https://doi.org/10.1080/07418825.2015.1064987

Wehby, J. H., & Kern, L. (2014). Intensive behavior intervention: What is it, what is its evidence base, and why do we need to implement now? *Teaching Exceptional Children*, *46*(4), 38–44. https://doi.org/10.1177/0040059914523956

Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. (2001). *Woodcock-Johnson III NU tests of achievement*. Riverside.

Younger, M., Warrington, M., & Williams, J. (1999). The gender gap and classroom interactions: Reality and rhetoric? *British Journal of Sociology of Education*, *20*(3), 325–341. https://doi.org/10.1080/01425699995290

Zucker, S. H., & Prieto, A. G. (1977). Ethnicity and teacher bias in educational decisions. *Journal of Instructional Psychology*, *4*(3), 2–5.