

Teacher Praise-to-Reprimand Ratios: Behavioral Response of Students at Risk for EBD Compared with Typically Developing Peers

Paul Caldarella

Ross A. A. Larsen

Leslie Williams

Brigham Young University

Howard P. Wills

University of Kansas

Joseph H. Wehby

Vanderbilt University

Abstract

Students with emotional and behavioral disorders (EBD) often receive low rates of teacher praise and high rates of teacher reprimands, though they may respond more positively to praise and more negatively to reprimands than their typically developing peers. Teacher praise-to-reprimand ratios (PRRs) are associated with increased student engagement, but more research is necessary to further explore how these ratios affect student behavior, particularly for students with or at risk for EBD. Commonly 3:1 or 4:1 is recommended as a desirable PRR, but students with EBD may need higher ratios to improve their classroom behavior. In our study of 540 students and 149 teachers, we found that as teacher PRR increased the engagement of at-risk students increased, though engagement for typically developing peers did not. A PRR of approximately 9:1 was needed for students at risk for EBD to

Author note: The research reported in this article was supported in part by a grant from the Institute of Education Sciences and the U.S. Department of Education (R324A120344), awarded to the University of Kansas in cooperation with Brigham Young University and Vanderbilt University. The opinions presented in this article are those of the authors, and no endorsement by the agency is intended or implied.

Address correspondence to: Paul Caldarella, PhD, Center for the Improvement of Teacher Education and Schooling, Brigham Young University, 149 MCKB, Provo, UT 84602. E-mail: paul_caldarella@byu.edu

approximate the engagement levels of their typically developing peers. No effect of PRR on disruption rates was found for either student category.

Keywords: praise, reprimands, ratio, elementary school, emotional and behavioral disorders

Students with emotional and behavioral disorders (EBD) experience challenges that include significantly lower perceptions of school climate and significantly higher rates of mental health problems and peer victimization than students without EBD (Salle, George, McCoach, Polk, & Evanovich, 2018). These students also show high rates of antisocial behaviors, internalizing symptoms, and attentional problems (Conley, Marchant, & Caldarella, 2014). They are more likely than their peers to experience major deficits in academic achievement across content areas (Nelson, Benner, Lane, & Smith, 2004; Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005), including writing and underlying reading performance significantly below their peers (Gage, Wilson, & MacSuga-Gage, 2014).

The negative academic outcomes for students with EBD may be related to teachers' struggles to manage these students' classroom behavior. For example, for boys with EBD the quality of teachers' classroom management in the first few years of school positively correlated with scores on standardized measures of third grade reading achievement (Garwood, Vernon-Feagans, & Family Life Project Key Investigators, 2017). Although students with EBD appear to become more engaged and less disruptive when they receive frequent teacher praise and infrequent reprimands (Downs et al., 2018), these students usually receive low rates of teacher praise and high rates of reprimands (Rathel, Drasgow, Brown, & Marshall, 2014; Sutherland & Wehby, 2001). The powerful impact of teacher praise on students with EBD may be explained, in part, by the limited amount of praise they receive from teachers (Rathel et al., 2014), making the praise they do receive particularly salient and effective (Downs et al., 2018). Teachers who deliver high rates of positive feedback compared to negative feedback may be particularly effective with students with EBD when they also provide multiple teaching and learning opportunities that enhance students' engagement (Rathel et al., 2014).

Praise-to-Reprimand Ratios

Of increasing interest for promoting school success for students with EBD is the study of teachers' praise-to-reprimand ratios (PRRs). Higher proportions of praise to reprimands have been associated with

lower levels of student noncompliance and higher levels of student enthusiasm and focus (Leff et al., 2011). In addition, students experiencing high PRRs have improved in their prosocial behavior, while students experiencing low PRRs have tended to decrease in their emotional regulation and concentration and to exhibit more disruptive behaviors (Reinke, Herman, & Newcomer, 2016). Unfortunately, reprimand rates in classrooms are commonly higher than praise rates (Reinke, Herman, & Stormont, 2013), as teachers often have difficulty maintaining high praise rates (Floress, Jenkins, Reinke, & McKown, 2018).

Classroom training and intervention can positively influence PRR. For example, implementing a group contingency program called Class-Wide Function-Related Intervention Teams (CW-FIT; Wills et al., 2010) has increased PRR in preschool (Jolstead et al., 2017), in general education kindergarten through second-grade classrooms (Caldarella, Williams, Hansen, & Wills, 2015), in third- and fifth-grade art classrooms (Nelson et al., 2018), and in a sixth-grade music classroom (Caldarella, Williams, Jolstead, & Wills, 2017). Increased PRRs were associated with increased on-task behaviors by students in all of these studies, but additional research is necessary to learn more about how these ratios affect student behavior, especially for students with or at risk for EBD. These students are commonly placed in inclusive general education classrooms for large portions of their school day (Floress et al., 2018) with teachers who have little knowledge and experience with EBD (Allday et al., 2012).

Suggestions for optimal PRRs have varied from 3:1 (Shores, Gunter, & Jack, 1993), to 4:1 (Trussell, 2008), to 5:1 (Flora, 2000). A 3:1 ratio is often recommended for general education classrooms, but students with EBD may need higher ratios to improve their behavior (Rathel et al., 2014). As noted in previous work (Downs et al., 2018; Sabey, Charlton, & Charlton, 2018), the most pressing questions that need to be addressed include how PRRs affect students with EBD, whether there is an optimal PRR for improving students' classroom behavior, and whether effects differ for students not identified with or at risk for EBD.

Study Purpose

The purpose of the current study was to examine the relationship between teacher PRRs and various behaviors of students who were and were not at risk for EBD in elementary school classrooms. As recommended by Downs and colleagues (2018), we also examined whether

there was an optimal PRR for students at risk for EBD compared to students not at risk. Three specific research questions were examined:

1. How are teacher PRRs related to the classroom engagement of students at risk for EBD compared to the engagement of typically developing peers?
2. How are teacher PRRs related to the classroom disruption rates of students at risk for EBD compared to the disruption rates exhibited by typically developing peers?
3. Is there an optimal PRR for improving the classroom behavior of students at risk for EBD?

Method

Setting and Participants

Over a three-year period, teacher and student data were gathered from participants in 19 elementary schools across Missouri, Tennessee, and Utah as part of a multisite randomized control trial (RCT) of CW-FIT (Wills, Wehby, Caldarella, Kamps, & Swinburne Romine, 2018). Students receiving free/reduced-price lunch (FRL) ranged from 35% to 99% ($M = 74.57$, $SD = 19.44$) across these schools. The current study included the data for control classroom participants from the study by Downs and colleagues (2018) as well as the treatment classroom participants from the RCT.

Teachers. Researchers met with district personnel to identify schools as potential sites for the RCT. Building administrators allowed researchers to present to their faculty to solicit teacher participants, who then completed informed consent materials required by institutional review boards (IRBs) at their respective universities and school districts. The majority of the 149 participating teachers were White/Caucasian (83%), female (95%), and had master's (49%) or bachelor's degrees (46%). Teachers ranged from 21 to 65 years old ($M = 38.51$, $SD = 12.51$), with teaching experience ranging from 0 to 44 years ($M = 9.20$, $SD = 9.01$). Data were collected during classroom instruction in the following subject areas: language arts (58%), math (37%), social studies (2%), science (1%), and other (2%). Both general education (93%) and special education (7%) settings were included. Teachers were randomly selected to be in treatment (52%) or control (48%) conditions, as described in the Procedures section.

Students. Participating teachers identified 311 students at risk for EBD and selected 229 as peers for comparison. See Table 1 for demographic information for these 540 students.

Table 1
Demographic Data for At-Risk Students and Peer-Comparison Students

Variable	<i>At-risk students</i> (<i>n</i> = 311)		<i>Peer-comparison students</i> (<i>n</i> = 229)	
	Frequency	Percentage	Frequency	Percentage
Gender				
Male	225	72.35	116	50.66
Female	86	27.65	101	44.10
Information missing	0	0	12	5.24
Grade				
K	55	17.68	38	16.59
1	58	18.65	42	18.34
2	50	16.08	37	16.16
3	65	20.90	49	21.40
4	35	11.25	29	12.66
5	36	11.58	25	10.92
6	11	3.54	7	3.06
Information missing	1	0.32	2	0.87
Ethnicity				
Black/African American	124	39.87	45	19.65
Hispanic/Latino	43	13.83	28	12.23
White/Caucasian	133	42.76	119	51.97
Asian/Pacific Islander	1	0.32	11	4.80
Other	4	1.29	2	0.87
Information missing	6	1.93	24	10.48
Primary language				
English	232	74.60	154	67.25
Other	14	4.50	12	5.24
Information missing	65	20.90	63	27.51
Disability				
No disability	243	78.14	180	78.60
SLD	24	7.72	5	2.18
SLI	3	0.96	3	1.31
DD	9	2.89	1	0.44
OHI	5	1.61	0	0
ASD	6	1.93	0	0
ED	4	1.29	0	0

(continued)

Table 1 (continued)
Demographic Data for At-Risk Students and Peer-Comparison Students

Variable	<i>At-risk students</i> (<i>n</i> = 311)		<i>Peer-comparison students</i> (<i>n</i> = 229)	
	Frequency	Percentage	Frequency	Percentage
ID	2	0.64	0	0
VI	1	0.32	0	0
Information missing	14	4.50	40	17.47
CW-FIT				
Treatment	170	54.66	120	52.40
Control	141	45.34	109	47.60

Note. SLD = specific learning disability, SLI = speech language impairment, DD = developmental disability, OHI = other health impairment, ASD = autism spectrum disorder, ED = emotional disturbance, ID = intellectual disability, and VI = visual impairment.

Identification of students at risk for EBD. Researchers used a gated multistep process to identify students at risk for EBD. Beginning with Stage 1 of the Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1992), a nationally normed systematic screening measure, participating teachers (a) studied definitions of internalizing and externalizing behaviors, (b) considered all students in their class and identified those who exhibited either form of behavior, and (c) rank ordered the selected students based on severity of behaviors. As a second step, teachers then completed the Problem Behavior scale (including both internalizing and externalizing subscales) of the Social Skills Improvement System (SSIS): Teacher Form (Gresham & Elliott, 2008) for those identified on the SSBD. The SSIS is a norm-referenced standardized measure used to identify students' strengths and weaknesses with three scales: Social Skills, Problem Behaviors, and Academic Competence.

Parents of students with scores in the *at risk* or higher range of the Problem Behavior scale completed informed consent/assent forms as outlined by the respective IRBs (typically two to three students per teacher). To confirm SSBD and SSIS results, researchers then directly observed those consented students using the Multi-Option Observational System for Experimental Studies (MOOSES; Tapp, Wehby, & Ellis, 1995). Students who exhibited disruptive behavior (minimum of 10 disruptions per 15 minutes) and/or low engagement levels (less than 75%) for at least two out of five observations met study criteria

to be included as students at risk for EBD, similar to criteria used in other studies (Kamps et al., 2011; Wills et al., 2018).

Selection of typically developing peers. When teachers had completed the SSBD, they also identified two or three students per class who exhibited cooperative and appropriate classroom behavior to serve as typically developing peer comparisons. Though teachers tried to gender match the peer-comparison students to the at-risk students, they did not do so in all cases. Peer-comparison students and their parents completed informed consent/assent forms as outlined by the respective IRBs, after which researchers collected direct observation data to compare with data from the students identified as at risk for EBD.

Direct Observation Process

Observers consisted of trained undergraduates (from a variety of majors) and graduate students (in education) supervised by master's level and doctoral level research staff, who also completed observations. Observers conducted 15-min direct observations on consented students using the MOOSES handheld computer software program in both treatment and control conditions spread over a period of four to six months. MOOSES enabled researchers to record behavior using frequency (i.e., student disruptive behaviors, teacher praise statement, and teacher reprimands) and duration (i.e., student engagement) codes for later analysis. See Table 2 for definitions and examples of frequency and duration codes.

In preparation, observers studied the MOOSES definitions for frequency and duration codes, then coded videos of students' behavior previously scored by researchers. When the videos were successfully coded at 85% accuracy, observers practiced observing student behavior in live classrooms with a research coordinator until they achieved 85% accuracy. Once proficiency was reached, data collection began.

In each classroom, researchers conducted observations during the subject that the teacher chose as most challenging to manage, at approximately the same time each day. Students identified as at risk and their teachers were observed on average 14.63 times ($SD = 2.71$; research goal 15 observations). Peers selected for comparison and their teachers were observed on average 3.17 times ($SD = 1.04$; research goal three observations), as the RCT emphasized improving the behavior of students at risk, and peers were observed as a norm

Table 2
Definition of Teacher and Student Variables

Variable	Definition	Examples	Non-examples
Teacher praise (frequency, event)	Verbal statements of approval going beyond an evaluation of adequacy or acknowledgement of a correct response to a question, directed to a target student or the group including the target student	"Good job of using indoor voices to work together, Table 3." Requests for high five, pat on back	"I've got Johnny's paper." Nonverbal praise, such as tokens or tickets for good behavior
Teacher reprimand (frequency, event)	Negative verbal comments such as scolding, threatening, or otherwise indicating disapproval with social behavior, or making comments with intent to correct behavior as it is occurring or after it has occurred; recorded if directed to the target student or a group of students including the target student	"Josh, quit wasting time and get back to work." "Start paying attention or your name is going on the board."	"This answer is incorrect." Nonverbal reprimands such as frowns, gestures, eyebrow raises, or tickets
Student disruptive behaviors (frequency, event)	Deliberate physical, motor, or verbal display of inappropriate behavior that interferes with a student's own participation and/or the productivity of peers; recorded as one occurrence unless the topography has changed, the behavior has ceased for at least 3 seconds, or teacher/student response separates the events	Rocking in chair Tapping/ drumming Throwing materials Laughing at peers Chatting that is not task related or teacher permitted Verbally refusing to comply	Borrowing a pencil or something related to the assignment and getting started on work right away Kneeling on chair to reach table that is difficult to reach when sitting

(continued)

Table 2 (continued)
Definition of Teacher and Student Variables

Variable	Definition	Examples	Non-examples
Student engagement (duration, event)	Student is appropriately participating, listening, or working on the assigned/approved activity.	Reading/writing Responding to questions Participating in groups/centers Listening/ watching Attending to material	Not responding when teacher gives a general request for information or answer to a question

Note. Target student = At-risk or peer-comparison student being observed with MOOSES.

reference. Researchers later calculated PRR using the following formula: *Total Teacher Praise Statements / (Total Teacher Praise Statements + Total Teacher Reprimands)*.

A reliability observer was present during 23.22% of observations. As calculated using MOOSES, an agreement was scored if within a 5-second window the same code was recorded (see Tapp et al., 1995). Interobserver agreement (IOA) was calculated as the number of agreements divided by the sum of the number of agreements plus disagreements. IOA for frequency codes was $M = 90.06\%$ ($SD = 14.08$) and for duration codes was $M = 99.21\%$ ($SD = 13.85$).

Procedures

Teachers and their classes were stratified by grade (K–2 and 3–6) and classroom type (e.g., general education or special education), and randomized to treatment or control condition using a selection function in Microsoft Excel. Students themselves were not randomized; rather their classes were randomly assigned. Teachers of both treatment and control classes identified the subject in their daily schedule that involved the most behavior problems for researchers to observe. After randomization, teachers in the treatment condition were trained on CW-FIT procedures and began implementing the intervention, which became their primary classroom management tool. Four main CW-FIT components were included: (a) direct social skills instruction, (b) interdependent group contingency, (c) positive reinforcement for

use of expected social skills, and (d) a secondary tier of support (e.g., self-management or help cards) for students needing additional intervention (see Wills et al., 2018, for more specific information on training and intervention procedures). Teachers of the control classrooms continued their routine management practices (e.g., token economies, class rewards, daily behavior report cards, praise, reprimands, loss of privileges, or clip charts).

Specific procedures were used in the RCT to prevent contamination. First, fidelity was monitored in control and treatment classes via direct observation. Data from observations confirmed that no significant treatment procedures were used in control classes during the study (see Wills et al., 2018, for details). Second, teachers signed a form agreeing they would not implement the treatment procedures if they were randomly selected for the control group. Finally, teachers in the control group were offered CW-FIT training, which was identical to treatment teachers in the spring of the school year following data collection.

Data Analysis Plan

We began with preliminary descriptive analyses to understand the nature of the data, then examined differences between at-risk and peer-comparison students across demographic and target variables. We created two multivariate multilevel regressions to answer the research questions regarding the relationships between PRR and student behaviors of engagement and disruptions.

Nested data required use of a clustering variable (student identification number) in the analysis. We used a multilevel analysis with time nested within student and student nested within teacher. We considered multivariate multilevel regression assumptions of linearity, independence, normality, variance equality, and multicollinearity. Independence was violated, as time was nested within students and students were nested within teacher. Residual plots showed no departure from normality or linearity.

Preliminary analyses. We first used independent samples *t*-tests to explore potential differences between at-risk students and comparison students. We then created a correlation table for the variables of interest.

Multivariate multilevel regression. As these data were from a RCT, causal relations can be appropriately inferred. We investigated variables of interest influenced by the intervention by using mediation analysis to suggest how the intervention affected the outcomes.

We ran a series of multivariate multilevel regressions in Mplus 8.1 (Muthén & Muthén, 2017) to answer the research questions. Due to convergence problems and less stringent assumptions, we used a Bayesian estimator for the parameter estimates. Initial models included covariates listed below, as well as grade level, site, class subject, praise rate, reprimand rate, PRR, curvilinear effects of PRR (PRR^2), and interactions with student at-risk designation (to investigate whether there was an optimal PRR). Covariates of the model included student status (at-risk = 1, peer = 0), student gender (male = 1, female = 0), student ethnicity (White = 1, other = 0), student disability status (no disability = 1, other = 0), CW-FIT status (treatment = 1, control = 0), and an interaction term (PRR with student at-risk status). A final model was also proposed, for which nonsignificant curvilinear and interaction effects had been excluded, to obtain an accurate p -value for the main effects.

Instrumental Variables Estimation

As noted above, data were from a RCT of CW-FIT (Wills et al., 2018), which attempted to increase the number of praise statements and thus the PRR. In a RCT, the process of randomization minimizes threats to internal validity that may confound casual statements. Thus because of randomization, other ways of explaining differences between the treatment and control groups on the outcomes of interest after intervention are minimized. Researchers used instrumental variables estimation (IVE; see Cook, Campbell, & Shadish, 2002), as it is a powerful tool in such cases to prove causal evidence of intermediate mechanisms.

An IVE must begin with a source of exogenous variation (i.e., imposed by outside forces) called an *instrument* (CW-FIT) and a *mediator* (PRR). Next, the relationship between the instrument and the outcomes (i.e., student engagement and student disruptions) needs to be completely mediated by the mediator. If the relationship between CW-FIT and student engagement and disruptions is completely mediated by PRR, then we can have a causal claim of PRR on the outcomes through calculating the indirect effect of CW-FIT on the student outcomes rather than simply correlating PRR with the outcomes. In this case, the model captures the effect that compliance to CW-FIT has on PRR and thus the student outcomes. As a note, in IVE *compliers* is a term referring to those participants in the experiment who comply with their treatment assignment (i.e., if assigned to treatment, they implement the intervention; if assigned to control, they do not).

Table 3
Descriptive Statistics by Student Group on Outcomes of Interest

Variables	At-risk students (<i>n</i> = 311)			Peer-comparison students (<i>n</i> = 229)			<i>t</i> ^a	<i>p</i>
	<i>M</i>	<i>SD</i>	min-max	<i>M</i>	<i>SD</i>	min-max		
Disruption rate	10.08	6.53	0.73–51.82	2.80	3.78	0.00–25.53	16.30	0.00
Engagement (%)	77.38	12.77	26.82–96.83	93.44	8.12	55.67–100.00	-17.82	0.00
Praise rate	2.19	1.78	0.00–10.56	2.36	2.29	0.00–9.50	-0.90	0.37
Reprimand rate	1.59	1.01	0.07–4.91	0.59	0.64	0.00–3.00	14.12	0.00
PRR	0.52	0.26	0.00–0.97	0.70	0.34	0.00–1.00	-6.57	0.00

Note. Data were averaged for all 15-minute observation periods. PRR = praise-to-reprimand ratio.

^aWelch's *t*-test was used assuming non-equal variances.

Results

As shown in Table 3, peer-comparison students had lower disruption rates, higher engagement, lower reprimands, and higher PRR than did the at-risk students (all $p < .05$). Praise rates were approximately equal between these two groups ($p = 0.37$).

As shown in Table 4, both PRR and praise rate were positively correlated with engagement, while reprimand rate and engagement were negatively associated. Teacher reprimand rate was associated with student disruptions, and PRR was negatively associated with student disruptions, though there was no significant correlation between praise rate and disruptions. CW-FIT was negatively associated with disruptions and positively associated with engagement and PRR (all $p < .05$). CW-FIT was not significantly associated with reprimand rate.

Results of the multivariate multilevel regression with PRR² and its interaction with students at risk was nonsignificant and was not shown in Table 4. The p -values for the curvilinear term and its interactions were always high, indicating the absence of any nonlinear trends in the data for both at-risk and peer-comparison students.

Table 4
Descriptive Statistics and Correlations of Variables of Interest Used
in Model at the Student Level (n = 540)

Variable	1	2	3	4	5	6	7
1. Disruption rate	1						
2. Engagement (%)	-0.64**	1					
3. Praise rate	-0.08	0.27**	1				
4. Reprimand rate	0.56**	-0.53**	-0.04	1			
5. PRR	-0.32**	0.51**	0.63**	-0.58**	1		
6. Students at risk	0.55**	-0.59**	-0.04	0.50**	-0.28**	1	
7. CW-FIT	-0.11**	0.23**	0.71**	-0.07	0.58**	0.01	1

** $p < .01$; CW-FIT = Class-wide function-related intervention teams; PRR = praise-to-reprimand ratio.

We ran a final model without the nonsignificant curvilinear terms to obtain trustworthy p -values (see Table 5). There was an interaction between PRR and students' risk status on student engagement ($p < .05$) but not on disruption rate ($p > .05$). A separate model (not shown) without the interaction term was run to test the main effect of PRR on disruption rate; the effect was nonsignificant ($p > .05$), though teacher reprimand rate was predictive of student disruption rate ($p < .05$).

Figure 1 shows the effect of PRR on engagement rates for both at-risk and peer-comparison students. Students who were at risk had a linear relation between PRR and engagement rate ($p < .05$), indicating that as the PRR increased so did engagement, though peers selected for comparison had no PPR-engagement relationship.

As shown in Figure 2 and Table 3, CW-FIT strongly predicted PRR, which in turn predicted student engagement rates. Figure 2 shows an IVE where treatment status was exogenous, as it was randomly assigned and thus was an instrumental variable. Figure 2 shows an unbiased treatment effect of CW-FIT ($\beta = 0.31$) for the compliers, assuming PRR was one of the mechanisms affecting outcomes. The main effect of CW-FIT on engagement was not significant in the presence of PRR, which is one of the assumptions of IVE (no third path).

A post-hoc exploratory analysis (not shown) was conducted interacting PRR with ethnicity to determine if PRR affected outcomes differently for minority students. No interaction effect was found ($p > .05$), indicating that the effects of PRR were the same regardless of students' minority status.

Table 5
Multivariate Multilevel Regression Results (n = 540 students)
Using a Bayesian Estimator of Teacher Praise-to-Reprimand Ratio
on Student Engagement and Student Disruptions

Predictor	Outcomes			
	Disruption rate		Engagement (%)	
	B(SD)	ES	B(SD)	ES
Praise rate	0.24(0.31)	0.07	0.57(0.63)	0.08
Reprimand rate	3.21(0.35)*	0.47	-4.03(0.73)*	0.26
Students at risk	2.88(0.86)*	0.25	-15.41(1.82)*	-0.58
PRR	1.90(3.00)	0.08	1.49(5.87)	0.03
PRR *Students at risk	0.26(2.05)	0.01	14.39(3.81)*	0.28
Female	-1.74(0.52)*	-0.15	3.32(0.86)*	0.12
No disability	1.34(.69)	0.08	1.52(1.28)	0.04
White student	-1.80(0.56)*	0.15	-0.40(1.07)	0.01
Grade level	-0.46(0.15)*	-0.14	-0.31(0.27)	-0.04
CW-FIT treatment	-2.54(0.99)*	0.22	-0.33(1.98)	-0.01
R ²	0.45(0.05)*	—	0.67(0.03)*	—

Note. All SD values are the posterior standard deviation (similar to standard error). ES = effect size (standardized betas) as reported in Mplus 8.1 output. PRR = praise-to-reprimand ratio. CW-FIT = class-wide function-related intervention teams. Controls (not shown) for within time effects have the following covariates: (a) daily praise rate, (b) daily reprimand rate, (c) teacher's training in CW-FIT treatment, (d) time, (e) daily PRR, and (f) site effects. * $p < .05$.

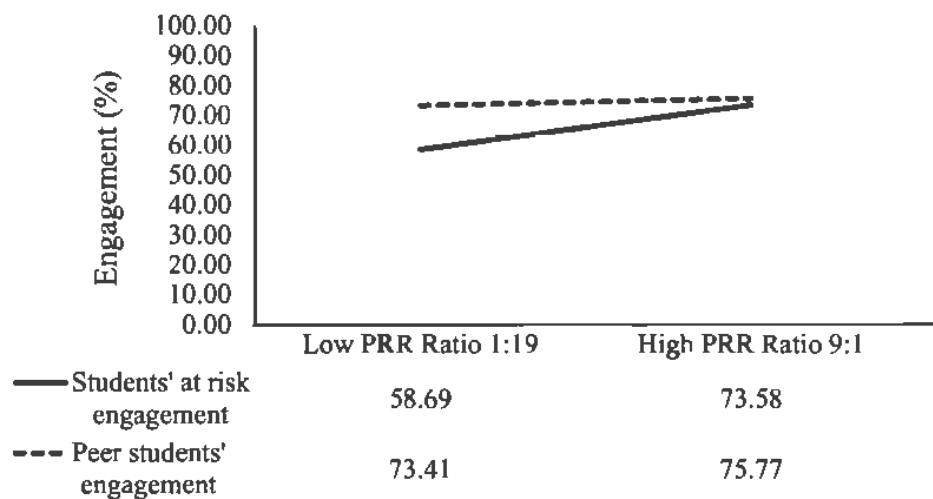


Figure 1. Interaction of student risk status and teacher praise-to-reprimand ratio (PRR) on student engagement.

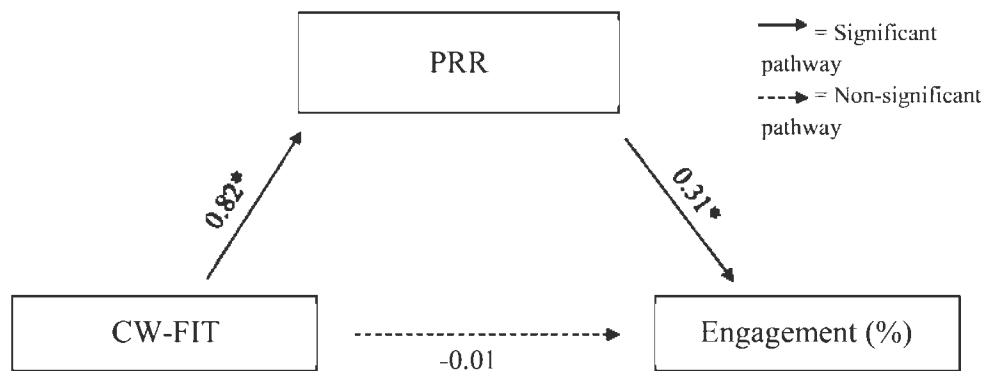


Figure 2. Mediation model to show how CW-FIT affects PRR for students at risk for EBD. CW-FIT = class-wide function-related intervention teams; PRR = praise-to-reprimand ratio. * $p < 0.05$.

Discussion

Students with EBD experience serious negative school outcomes, including lower perceptions of school climate, significantly higher rates of mental health problems and peer victimization, and lower academic achievement than students without EBD (Gage et al. 2014; Nelson et al., 2004; Salle et al., 2018; Wagner et al., 2005). The results of this study regarding ways teacher behaviors may help mitigate such outcomes can be critical to the school experience of these students. The quality of teachers' classroom management appears to influence the academic success of students with EBD (Garwood et al., 2017) as well as their social and emotional experience.

Answers to the Research Questions

The first research question examined how teacher PRRs relate to the classroom engagement of students at risk for EBD when compared to the engagement of peers selected for comparison. Results showed no effect of PRR on engagement for peer-comparison students, though this finding may have resulted from a ceiling effect because peer-comparison students had high engagement rates throughout the study. However, for students at risk for EBD a linear relationship between PRR and engagement rate indicated that increasing PRR increased engagement, similar to results of research by others (Leff et al., 2011). As in the research of Downs et al. (2018), at-risk students showed more sensitivity to teacher PRR than did their more typical

peers. However, ironically students with EBD typically receive less teacher praise and more teacher reprimands (Rathel et al., 2014; Sutherland & Wehby, 2001), unless intervention is applied. Unfortunately, teachers find it difficult to maintain high rates of classroom praise (Floress et al., 2018). IVE analysis confirmed that the effect of PRR on engagement was via the CW-FIT intervention.

The second research question examined how teacher PRRs related to the classroom disruption rates of students at risk for EBD compared to rates of peers not at risk. PRR was negatively correlated with student disruptions, but the directionality of the effect was not clear, producing results similar to those of Leff et al. (2011). Results from the multivariate multilevel regressions showed no causal effect of PRR on disruption rates for either category of students. The bivariate correlation does not isolate the causal effect of PRR on disruptions, which helps explain the differences in the multivariate multilevel regression results.

In the current study, reprimand rates predicted disruption rates for all students. More teacher reprimands may be directed to a student who is more disruptive, but there is also a possibility that a student is more disruptive because he or she receives more reprimands. Additional research is needed on this issue to identify possible causal mechanisms. As reprimands appear to decrease engagement and increase disruptions, we recommend focusing on increasing praise and PRR and avoiding reprimands whenever possible.

The last research question examined whether there is an optimal PRR for improving the classroom behavior (i.e., engagement and disruptions) of students at risk for EBD. Results showed an absence of nonlinear trends in the data for either the student at-risk group or the peer-comparison students, suggesting that there was no *tipping point* for the effect of PRR on engagement and disruption rates. Though we did not find a tipping point at which the behavior of students at risk improved dramatically, we found a linear relationship indicating that the higher the PRR for the students at risk, the higher their engagement. Thus, students at risk for EBD appear to benefit from any increase in the PRR regardless of magnitude. However, if the goal is to have equal engagement among all students, results suggest an estimated 9:1 PRR would be required (see Figure 1). For this study the CW-FIT intervention was necessary to get to this 9:1 ratio, which differs from PRR recommendations made by others of 3:1 (Shores et al., 1993), 4:1 (Trussell, 2008), or 5:1 (Flora, 2000).

Results also showed that student gender (male) and minority status predicted higher disruption rates, while only gender (female) predicted higher engagement (see Table 5). This could be explained by

the higher rates of male and minority students in the student at-risk category compared to the peer comparison group. Results also showed that the effects of PRR were the same regardless of minority status.

Limitations and Future Research

One notable limitation of this study was the disproportion between African American students designated as at risk and those selected for peer comparison, which may represent issues of disproportionality within risk identification (Osher et al., 2004). Though we had a relatively diverse sample of students identified as at risk for EBD, few of them had been classified by their schools as having EBD. A study including more students actually classified with EBD, with stricter inclusion criteria (e.g., students with an individualized education program), is warranted. The large number of students found in the current study to be at risk but not identified as having EBD and receiving special education services is concerning but is similar to what has been found by other researchers (Wagner et al., 2005). As many states are utilizing a response to intervention approach for special education classification, educators may find it beneficial to identify students as at risk and to provide proactive classroom interventions such as CW-FIT before characteristics of EBD become more severe.

Another limitation is that we did not identify types of teacher praise. Future studies could examine the potentially differential effects of behavior specific versus general praise, as teachers' behavior specific praise has been found to be associated with decreased student off-task behavior (Floress et al., 2018; Sabey et al., 2018).

Some limitations were related to students selected for peer comparison. The identification process for peer-comparison students was less systematic than the screening for at-risk students, leading to a question of whether the peer-comparison students accurately represented average elementary school students. In addition, the peer comparison students were observed less frequently than at-risk students, perhaps resulting in a less precise view of their behavior. Nevertheless, the peer-comparison students exhibited characteristics typically expected of elementary students (e.g., high engagement rates and low disruption rates), leading the researchers to be comfortable with this generalization. Future researchers could use other methods to select students for comparison, such as propensity score matching (Murnane & Willett, 2010).

We recommend several related areas for future research. The effects of PRR on student grades and test scores warrant attention.

We acknowledge that our 9:1 PRR recommendation for students at risk for EBD is an estimate of the actual ratio needed; future work could further explore this ratio. The causal direction between the reprimand rate and disruption rate also remains unclear; future studies would be helpful in isolating the directionality of this relationship.

Beneficial future research could extend this study by examining the effects of PRR on student engagement and disruption rates in middle schools and high schools. PRR has been shown to worsen (increasingly more reprimands than praise) as students' grade levels increase (Reddy, Fabiano, Dudek, & Hsu, 2013).

Conclusion and Implications

Results suggest that PRR is important for increasing the engagement of students who are at risk for EBD. Though there does not appear to be a "tipping point" at which student behavior drastically improves, as teacher PRR increases so does student engagement. However, a high ratio (9:1) appears to be needed to help at-risk students approximate the engagement levels of their more typical peers. A positive behavior intervention, such as CW-FIT, can enable teachers to achieve such a ratio. This study contributes substantially, as students with EBD are commonly placed in inclusive classrooms for much of their school day (Floress et al., 2018), with teachers who have little training in how to work with such students (Allday et al., 2012). Teachers may utilize study results to improve classroom environments by increasing their PRR and, thus, improve the behavior of elementary students with EBD risk. Teachers should be trained and supported in giving high rates of praise and low rates of reprimands to improve student engagement. While the study results are promising, additional research is needed examining the impact of teacher PRR on student behavior, particularly for students with or at risk for EBD.

References

- Allday, R. A., Hinkson-Lee, K., Hudson, T., Neilsen-Gatti, S., Kleinke, A., & Russel, C. S. (2012). Training general educators to increase behavior-specific praise: Effects on students with EBD. *Behavioral Disorders, 37*, 87–98. doi:10.1177/019874291203700203
- Caldarella, P., Williams, L., Hansen, B. D., & Wills, H. (2015). Managing student behavior with class-wide function-related intervention teams: An observational study in early elementary

- classrooms. *Early Childhood Education Journal*, 43, 357–365. doi:10.1007/s10643-014-0664-3
- Caldarella, P., Williams, L., Jolstead, K. A., & Wills, H. P. (2017). Managing student behavior in an elementary school music classroom: A study of class-wide function-related intervention teams. *Update: Applications of Research in Music Education*, 35(3), 23–30. doi:10.1177/8755123315626229
- Conley, L., Marchant, M., & Caldarella, P. (2014). A comparison of teacher perceptions and research-based categories of student behavior difficulties. *Education*, 134, 439–451.
- Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Downs, K. R., Caldarella, P., Larsen, R. A. A., Charlton, C. T., Wills, H. P., Kamps, D. M., & Wehby, J. H. (2018). Teacher praise and reprimands: The differential response of students at risk for emotional and behavioral disorders. *Journal of Positive Behavior Interventions*. Advance online publication. doi:10.1177/1098300718800824
- Flora, S. R. (2000). Praise's magic reinforcement ratio: Five to one gets the job done. *The Behavior Analyst Today*, 1, 64–69. doi:10.1037/h0099898
- Floress, M. T., Jenkins, L. N., Reinke, W. M., & McKown, L. (2018). General education teachers' natural rates of praise: A preliminary investigation. *Behavioral Disorders*, 43, 411–422. doi:10.1177/0198742917709472
- Gage, N. A., Wilson, J., & MacSuga-Gage, A. S. (2014). Writing performance of students with emotional and/or behavioral disabilities. *Behavioral Disorders*, 40, 3–14. doi:10.17988/0198-7429-40.1.3
- Garwood, J. D., Vernon-Feagans, L., & Family Life Project Key Investigators. (2017). Classroom management affects literacy development of students with emotional and behavioral disorders. *Exceptional Children*, 83, 123–142. doi:10.1177/0014402916651846
- Gresham, F. M., & Elliott, S. N. (2008). *Social Skills Improvement System: Rating scales*. Bloomington, MN: Pearson Assessments.
- Jolstead, K. A., Caldarella, P., Hansen, B., Korth, B. B., Williams, L., & Kamps, D. (2017). Implementing positive behavior support in preschools: An exploratory study of CW-FIT tier 1. *Journal of Positive Behavior Interventions*, 19, 48–60. doi:10.1177/1098300716653226

- Kamps, D., Wills, H. P., Heitzman-Powell, L., Laylin, J., Szoke, C., Petrillo, T., & Culey, A. (2011). Class-Wide Function-Related Intervention Teams: Effects of group contingency programs in urban classrooms. *Journal of Positive Behavior Interventions, 13*, 154–167. doi:10.1177/1098300711398935
- Leff, S. S., Thomas D. E., Shapiro E. S., Paskewich B., Wilson K., Necowitz-Hoffman, B., & Jawad, A. F. (2011). Developing and validating a new classroom climate observation assessment tool. *Journal of School Violence, 10*, 165–184. doi:10.1080/15388220.2010.539167
- Murnane, R. J., & Willett, J. B. (2010). *Methods matter: Improving causal inference in educational and social science research*. Oxford, UK: Oxford University Press.
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic achievement of K–12 students with emotional and behavioral disorders. *Exceptional Children, 71*, 59–73. doi:10.1177/001440290407100104
- Nelson, M. A., Caldarella, P., Hansen, B. D., Graham, M. A., Williams, L., & Wills, H. P. (2018). Improving student behavior in art classrooms: An exploratory study of CW-FIT. *Journal of Positive Behavior Interventions, 20*, 227–238. doi:10.1177/1098300718762744
- Osher D. A., Cartledge, G. W., Oswald, D. O., Sutherland, K. S., Artiles, A. J., & Coutinho, M. A. (2004). Cultural and linguistic competency and disproportionate representation. In R. B. Rutherford, M. M. Quinn, & S. R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 54–77). New York, NY: Guilford Press.
- Rathel, J. M., Drasgow, E., Brown, W. H., & Marshall, K. J. (2014). Increasing induction-level teachers' positive-to-negative communication ratio and use of behavior-specific praise through e-mailed performance feedback and its effect on students' task engagement. *Journal of Positive Behavior Interventions, 16*, 219–233. doi:10.1177/1098300713492856
- Reddy, L. A., Fabiano, G. A., Dudek, C. M., & Hsu, L. (2013). Instructional and behavior management practices implemented by elementary general education teachers. *Journal of School Psychology, 51*(6), 683–700. doi:10.1016/j.jsp.2013.10.001

- Reinke, W. M., Herman, K. C., & Newcomer, L. (2016). The Brief Student-Teacher Classroom Interaction Observation: Using dynamic indicators of behaviors in the classroom to predict outcomes and inform practice. *Assessment for Effective Intervention, 42*, 32–42. doi:10.1177/1534508416641605
- Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS: Identifying areas for enhancement. *Journal of Positive Behavior Interventions, 15*, 39–50. doi:10.1177/1098300712459079
- Sabey, C. V., Charlton, C., & Charlton, S. R. (2018). The “magic” positive-to-negative interaction ratio: Benefits, applications, cautions, and recommendations. *Journal of Emotional and Behavioral Disorders*. Advance online publication. doi:10.1177/1063426618763106
- Salle, T. L., George, H. P., McCoach, D. B., Polk, T., & Evanovich, L. L. (2018). An examination of school climate, victimization, and mental health problems among middle school students self-identifying with emotional and behavioral disorders. *Behavioral Disorders, 43*, 383–392. doi:10.1177/0198742918768045
- Shores, R. E., Gunter, P. L., & Jack, S. L. (1993). Classroom management strategies: Are they setting events for coercion? *Behavioral Disorders, 18*, 92–102. doi:10.1177/019874299301800207
- Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *The Journal of Special Education, 35*, 161–171. doi:10.1177/002246690103500306
- Tapp, J., Wehby, J., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instruments, & Computers, 27*, 25–31. doi:10.3758/BF03203616
- Trussell, R. P. (2008). Classroom universals to prevent problem behaviors. *Intervention in School and Clinic, 43*, 179–185. doi:10.1177/1053451207311678
- Wagner, M., Kutash, K., Duchnowski, A. J., Epstein, M. H., & Sumi, W. C. (2005). The children and youth we serve: A national picture of the characteristics of students with emotional disturbances receiving special education. *Journal of Emotional and Behavioral Disorders, 13*, 79–96. doi:10.1177/10634266050130020201
- Walker, H. M., & Severson, H. H. (1992). *Systematic Screening for Behavior Disorders* (2nd ed.). Longmont, CO: Sopris West.

- Wills, H. P., Kamps, D., Hansen, B. D., Conklin, C., Bellinger, S., Neaderhiser, J., & Nsubuga, B. (2010). The classwide function-based intervention team program. *Preventing School Failure: Alternative Education for Children and Youth, 54*, 164–171. doi:10.1080/10459880903496230
- Wills, H., Wehby, J., Caldarella, P., Kamps, D., & Swinburne Romine, R. (2018). Classroom management that works: A replication trial of the CW-FIT program. *Exceptional Children, 84*, 437–456. doi:10.1177/0014402918771321